
University
LECTURE
Series

Volume 4

Group Characters,
Symmetric Functions,
and the Hecke Algebras

David M. Goldschmidt



American Mathematical Society
Providence, Rhode Island

VISIT...

LANZAROTE
Caliente.COM

1991 *Mathematics Subject Classification*. Primary 20C15, 20C30,
20F36, 57M25.

Library of Congress Cataloging-in-Publication Data

Goldschmidt, David M.

Group characters, symmetric functions, and the Hecke algebra/David M. Goldschmidt.

p. cm. —(University lecture series, ISSN 1047-3998; v. 4)

Lectures given at Berkeley in spring of 1989.

Includes bibliographical references.

ISBN 0-8218-7003-3 (alk. paper)

1. Hecke algebras. 2. Representations of groups. 3. Finite groups. I. Title. II. Series:

University lecture series (Providence, R.I.); v. 4.

QA174.2.G65 1993

92-43284

512'.2—dc20

CIP

Copying and reprinting. Individual readers of this publication, and nonprofit libraries acting for them, are permitted to make fair use of the material, such as to copy a chapter for use in teaching or research. Permission is granted to quote brief passages from this publication in reviews, provided the customary acknowledgment of the source is given.

Republication, systematic copying, or multiple reproduction of any material in this publication (including abstracts) is permitted only under license from the American Mathematical Society. Requests for such permission should be addressed to the Manager of Editorial Services, American Mathematical Society, P.O. Box 6248, Providence, Rhode Island 02940-6248.

The owner consents to copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law, provided that a fee of \$1.00 plus \$.25 per page for each copy be paid directly to the Copyright Clearance Center, Inc., 27 Congress Street, Salem, Massachusetts 01970. When paying this fee please use the code 1047-3998/93 to refer to this publication. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale.

Copyright ©1993 by the American Mathematical Society. All rights reserved.

The American Mathematical Society retains all rights
except those granted to the United States Government.

Printed in the United States of America

The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability. ♾

This publication was typeset using $\mathcal{A}\mathcal{M}\mathcal{S}$ - $\text{\TeX}$,
the American Mathematical Society's $\text{\TeX}$ macro system.

10 9 8 7 6 5 4 3 2 01 00 99 98 97

Contents

Preface	vii
---------	-----

Part I

Chapter 1.	Finite-Dimensional Algebras	1
Chapter 2.	Group Characters	5
Chapter 3.	Divisibility	11
Chapter 4.	Induced Characters	15
Chapter 5.	Further Results	19

Part II

Chapter 6.	Permutations and Partitions	27
Chapter 7.	The Irreducible Characters of S^n	33
Chapter 8.	The Specht Modules	37
Chapter 9.	Symmetric Functions	43
Chapter 10.	The Schur Functions	47
Chapter 11.	The Littlewood-Richardson Ring	51
Chapter 12.	Two Useful Formulas	55

Part III

Chapter 13.	The Hecke Algebra	61
Chapter 14.	The Markov Trace	67
Bibliography		73

Preface

These are the lecture notes from a course I gave at Berkeley in the spring of 1989. My original motivation was to understand the formulae for the “weights” (i.e., Fourier coefficients) of the Markov trace which were computed by Ocneanu [3]. Luckily for me, halfway through the course Vaughan Jones showed me a slick preprint by Springer [9] which gave a very simple explanation of Ocneanu’s results, and I have incorporated this approach in §14. The prerequisites for this are §1 and most of Part II. But it is important to point out that one needs to know almost nothing in order to prove that the trace exists and to obtain a constructive algorithm for calculating it. The reader who is only interested in this aspect can just read (13.1), (13.2), (13.3), and (14.1), and should definitely see [5].

The course was organized into three parts, as is clear from the table of contents. In Part I, I prove Burnside’s Theorem that groups of order $p^a q^b$ are solvable (3.6), Frobenius’ Theorem on the existence of Frobenius kernels (4.5), and Brauer’s characterization of characters (5.8). As an application of the latter, I prove Brauer’s theorem on blocks of defect zero (5.11). The reader who is only interested in the later material can skip all of §5 and most, if not all, of §3. The most important results which are needed in Part II are the basic facts about induced characters, primarily Mackey’s theorem (4.4).

The material in Part II is far from complete, the most glaring omission being the Littlewood-Richardson rule. I first give an algorithm for computing the character table of S^n (§7) and I construct the Specht modules (§8) following James [4]. Following Macdonald [8], I next derive the “determinant form” (11.4) for the irreducible characters of S^n using the theory of symmetric functions, and I then obtain the hook-length formula (12.1) and the Murnaghan-Nakayama formula (12.6) as consequences.

In Part III, I prove that the field of rational functions is a splitting field for the Hecke algebra by first extending scalars to the field of formal Laurent series and then descending. The reader who is content to just use Laurent series can skip this and save a little time. I then develop Springer’s observation that the Fourier transform of the Markov trace is really a homomorphism from the ring of symmetric functions to the field of rational functions in two variables.

Part I

CHAPTER 1

Finite-Dimensional Algebras

In this section, all algebras will be finite-dimensional algebras with identity. Let F be a field of characteristic zero and let A be an F -algebra. For $a \in A$, let $a_R : A \rightarrow A$ be right multiplication by a ; then the map $a \mapsto a_R$ embeds A into $\text{End}_F(A)$ because A has an identity. There is a symmetric bilinear form on A called the *trace form* which is given by

$$(a, b) = \text{tr}(a_R b_R).$$

This form satisfies the important identity

$$(1.1) \quad (ax, b) = (a, xb) \quad \text{for all } a, b, x \in A.$$

When the trace form is nondegenerate, we will say that A is *semisimple*. Let $\{e_1, \dots, e_n\}$ be a basis of A ; then A is semisimple if and only if the *discriminant* $\Delta(A) = \det(e_i, e_j)$ is nonzero. This condition is invariant under extension of scalars, for if K is an extension field of F , then $\{e_1 \otimes 1, \dots, e_n \otimes 1\}$ is a basis of $A \otimes_F K$ over K and $(e_i \otimes 1, e_j \otimes 1) = (e_i, e_j)$.

More generally, (1.1) implies that the radical of the trace form, which we shall denote by $J(A)$, is a 2-sided ideal of A . If $x \in J(A)$, then $\text{tr}(x_R^n) = 0$ for all n which implies by standard linear algebra that x_R , and hence x itself, is nilpotent (characteristic zero is used here!). Conversely, it is clear that any right ideal consisting entirely of nilpotent elements must be contained in $J(A)$, so $J(A)$ is characterized as the largest right ideal of A consisting entirely of nilpotent elements (it is not hard to show that $J(A)$ is actually a nilpotent ideal). In particular, $J(A/J(A)) = 0$.

Let M be an irreducible A -module (i.e., M is nonzero and has no proper submodules) and let m be a nonzero element of M . Then $M = mA$, and the map $a \mapsto ma$ is an A -module epimorphism whose kernel, call it I , is a maximal right ideal of A . We claim that $J(A) \subseteq I$, for if not, then $A = I + J(A)$ and we can write $1 = x + y$ with $x \in I$ and $y \in J(A)$. But then $x = 1 - y$ is invertible since y is nilpotent, which is a contradiction. We have proved

(1.2) *Let A be an F -algebra and let $J(A)$ be the radical of the trace form. Then $J(A)$ is the largest right ideal of A consisting entirely of nilpotent elements, $J(A)$ annihilates every irreducible A -module, and $A/J(A)$ is semisimple. Moreover, $J(A) = 0$ iff $J(A \otimes_F K) = 0$ for every extension field K of F . $\square$*

We say that a ring is *simple* if it has no proper 2-sided ideals. Since $J(A)$ is a 2-sided ideal, simple rings are semisimple. For the remainder of this section, we assume that A is semi-simple.

Let I be a minimal nonzero 2-sided ideal of A and put $I' = I^\perp = \{x \in A \mid (y, x) = 0 \text{ for all } y \in I\}$. Then (1.1) implies that I' is also a 2-sided ideal. By the minimality of I , either $I \subseteq I'$ or $I \cap I' = 0$. If $I \subseteq I'$, we would have $\text{tr}(x_R^2) = 0$ for all $x \in I$ and then $I \subseteq J(A) = 0$ which is not the case. Therefore we must have $I \cap I' = 0$, whence it follows by linear algebra that $A = I \oplus I'$ and $II' = I'I = 0$.

Writing $1 = e + e'$ with $e \in I$ and $e' \in I'$, it is immediate that e and e' are central idempotents which act as 2-sided identities for I and I' respectively, and we conclude that the sum $A = I \oplus I'$ is an algebra direct sum. Moreover, $I = Ie$, $I' = I'e'$, and I and I' are both semisimple (by (1.2) or directly). By minimality, I is a simple algebra. By induction on $\dim(A)$, I' is an algebra direct sum of its minimal 2-sided ideals, each of which is a simple algebra.

A central idempotent e is called *imprimitive* if there exist nonzero central idempotents e_1, e_2 with $e = e_1 + e_2$ and $e_1 e_2 = 0$, otherwise, e is called *primitive*. From the above, we see that if $\{e_1, \dots, e_s\}$ is the set of primitive central idempotents of A , then $\{Ae_1, \dots, Ae_s\}$ is the set of minimal 2-sided ideals of A , $1 = \sum_{i=1}^s e_i$, and $A = \sum_i \bigoplus Ae_i$.

Put $B_i = Ae_i$ ($1 \leq i \leq s$) and let M be an irreducible A -module. Then $M = \sum_i \bigoplus Me_i$, so there is a unique i for which $M = Me_i$, and $Me_j = 0$ for $i \neq j$. Thus, M is an irreducible B_i -module. Since B_i is simple, M is faithful, i.e., $\{x \in B_i \mid Mx = 0\} = 0$. Let I be a minimal right ideal of B_i . Then MI is a nonzero submodule of M and hence we can choose $m \in M$ with $mI \neq 0$. Since mI is a submodule of M , $mI = M$ and the map $x \mapsto mx$ defines a nonzero homomorphism of irreducible B_i -modules $I \rightarrow M$ which is therefore an isomorphism. Summarizing our observations thus far, we have proved

(1.3) *Let A be a semisimple algebra, and let $\{e_1, \dots, e_s\}$ be the set of primitive central idempotents of A . Then $\{Ae_1, \dots, Ae_s\}$ is the set of minimal 2-sided ideals of A . Each Ae_i is a simple algebra with identity e_i and A is the algebra direct sum*

$$A = \sum_{i=1}^s \bigoplus Ae_i.$$

Moreover, if M_i is a minimal right ideal of Ae_i , then $\{M_1, \dots, M_s\}$ is a set of representatives for the isomorphism classes of irreducible A -modules.

Much more can be said about the structure of the simple components of A . Namely, we have

1.4 (Rieffel-Wedderburn). *Let B be a simple ring with identity, I a right ideal of B , and $D = \text{End}_B(I)$. Then the natural map $B \mapsto \text{End}_D(I)$ given by right multiplication is an isomorphism.*

PROOF. Since B is simple, $B = BI$. Now for $b \in B$, let b_R denote right multiplication by b , then $b_R \in \text{End}_D(I)$. Choose $x \in I$ and let $x_L : I \rightarrow I$ denote left multiplication by x . Notice that $x_L \in D$ since $x(yb) = (xy)b$ for all $y \in I$ and all $b \in B$. Thus for any $\varphi \in \text{End}_D(I)$ we have $\varphi(xy) = x\varphi(y)$ for all $x, y \in I$. In particular, for any $b \in B$ we get $\varphi(xby) = \varphi((xb)y) = xb\varphi(y)$. Letting x range over I while fixing b, y , and φ yields $\varphi \circ (by)_R = (b\varphi(y))_R$. This says that the image of $B = BI$ under the natural map is a left ideal of

$\text{End}_D(I)$. But the image of B contains 1, and is therefore equal to $\text{End}_D(I)$. Thus, the natural map is onto. Since B is simple, the result follows. $\square$

We remark that if I is a minimal right ideal above then D must be a division ring. For if $0 \neq \varphi \in D$, then $\ker(\varphi)$ and $\text{im}(\varphi)$ are both right ideals of A , whence $\ker(\varphi) = 0$ and $\text{im}(\varphi) = I$. (This observation is usually known as *Schur's Lemma*.) Since finite-dimensional algebras always have minimal right ideals, we have

(1.5) **COROLLARY.** *Let B be a simple finite-dimensional algebra, let I be a minimal right ideal of B , and let $D = \text{End}_B(I)$. Then D is a (finite-dimensional) division algebra, and B is isomorphic to the algebra of $n \times n$ matrices over D , where $n = \dim_D(I)$. $\square$*

We apply (1.5) with $B = B_i$ a simple component of A and $I = I_i$ a minimal right ideal of A contained in B . By (1.3) I is a minimal right ideal of B . By (1.5), $D = D_i = \text{End}_A(I_i)$ is a finite-dimensional division algebra over F , I is a D -vector space of dimension $n = n_i$, say, and $\text{End}_D(I)$ is the ring of $n \times n$ matrices over D .

Let e be the matrix with a 1 in position $(1, 1)$ and zeros elsewhere. Then eB is the right ideal of all matrices whose non-zero entries are in row 1. In particular, $\dim_D(eB) = n$ so eB is a minimal right ideal. Since $e^2 = e$ we have $B = eB \oplus (1 - e)B$, whence eB is an A -module direct summand of A and is therefore projective. By (1.3) we conclude that all irreducible A -modules are projective. By general nonsense (see e.g. [7]) we then have

(1.6) *Every A -module is a direct sum of irreducible A -modules.* $\square$

We next consider the effect of extending the field of scalars.

(1.7) *Let K be an extension field of F and let $A_K = A \otimes_F K$. Continuing the notation of (1.3), every irreducible A_K -module is a constituent of exactly one of the A_K -modules $M_i \otimes_F K$.*

PROOF. By (1.2) A_K is semisimple. The map $a \mapsto a \otimes 1$ is an embedding, and we will identify A with $A \otimes 1$. The primitive central idempotents e_i of A are still central idempotents of A_K , but they may no longer be primitive. Write $e_i = \sum_{j=1}^{m_i} e_{ij}$ where the e_{ij} are primitive central idempotents of A_K . Since two primitive central idempotents are either equal or orthogonal, and the e_i are orthogonal, it follows that $e_{ij}e_{kl} = \delta_{ik}\delta_{jl}$. In particular, the e_{ij} are distinct. Let

$$\{M_{ij} \mid 1 \leq i \leq s, 1 \leq j \leq m_i\}$$

be a corresponding set of minimal right ideals of A_K and let $k \neq i$. Then $M_{ij}e_k = M_{ij}e_{ij}e_k = 0$ which implies that M_{ij} is not a constituent of $M_k \otimes_F K$ for all j . But by (1.6), A is a module direct sum of minimal right ideals, so A_K is a corresponding direct sum of right ideals of the form $M_i \otimes_F K$. Since each M_{ij} is a constituent of A_K , we conclude that M_{ij} is a constituent of $M_i \otimes_F K$. $\square$

In the special case where $\text{End}_A(M) = F$, we say that the irreducible A -module M is *absolutely irreducible*. In this case, (1.5) says that the corresponding minimal 2-sided ideal is a complete matrix algebra over F . We say that an extension field K of F is a *splitting field* for A if every irreducible

A_K -module is absolutely irreducible. Splitting fields certainly exist, for example if K is algebraically closed then there are no nontrivial finite-dimensional division algebras over K (see, e.g., [7]), so K is necessarily a splitting field.

Here is a useful criterion for a module to be absolutely irreducible:

(1.8) *Let M be an A -module with $\dim_F(M) = n$, and suppose that the direct sum of i copies of M is a module direct summand of A . Then $i \leq n$ with equality possible iff M is absolutely irreducible.*

PROOF. By extending the field of scalars and changing notation if necessary, we may assume that F is a splitting field. Refine the direct sum decomposition of (1.3) to a direct sum of minimal right ideals using (1.6). Since the minimal 2-sided ideals of A are complete matrix algebras over F , an easy dimension count shows that each minimal right ideal occurs with multiplicity equal to its degree. (Note that minimal right ideals belonging to different 2-sided ideals are not isomorphic by (1.3).)

Let N be an irreducible constituent of M . Then it follows from the Jordan-Holder theorem that $i \leq \dim(N) \leq \dim(M)$ with equality iff $M = N$. $\square$

We conclude this section with a useful result on the product of semisimple algebras.

(1.9) *Suppose that A_1 and A_2 are semisimple F -algebras and let $A = A_1 \otimes_F A_2$. Then A is semisimple. If F is a splitting field for both A_1 and A_2 then an A -module M is irreducible iff $M \cong M_1 \otimes_F M_2$ where M_i is an irreducible A_i -module ($i = 1, 2$).*

PROOF. If $a_i \in A_i$ ($i = 1, 2$) then $(a_1 \otimes a_2)_R = a_{1R} \otimes a_{2R}$, $\text{tr}((a_1 \otimes a_2)_R) = \text{tr}(a_1) \text{tr}(a_2)$, and hence $(a_1 \otimes a_2, b_1 \otimes b_2) = (a_1, b_1)(a_2, b_2)$. This implies that $\Delta(A) = \Delta(A_1)\Delta(A_2) \neq 0$, so A is semisimple.

Now suppose that F is a splitting field for both A_1 and A_2 , and let I_i be a minimal right ideal of A_i of dimension n_i . Then by (1.8) the multiplicity of I_i in A_i is n_i ($i = 1, 2$). Since $\dim(I_1 \otimes I_2) = n_1 n_2$, which is the multiplicity of $I_1 \otimes I_2$ in $A_1 \otimes A_2$, we are done by (1.8). $\square$

CHAPTER 2

Group Characters

In this section, we apply the results of §1 to the *group algebra* of a finite group. First, we fix some standard notation:

G	A finite group
$\mathbf{C}$	The complex numbers
V	A finite-dimensional complex vector space
$\text{End}(V)$	The ring of linear transformations on V
$\mathbf{GL}(V)$	The group of invertible elements of $\text{End}(V)$
$M(n, \mathbf{C})$	The ring of $n \times n$ complex matrices
$\mathbf{GL}(n, \mathbf{C})$	The group of invertible elements of $M_n(\mathbf{C})$

A *linear representation* of G is a homomorphism $\mathcal{X} : G \rightarrow \mathbf{GL}(V)$ for some V . A *matrix representation* of G is a homomorphism $\mathcal{X} : G \rightarrow \mathbf{GL}(n, \mathbf{C})$ for some n . Two linear (resp. matrix) representations $\mathcal{X}, \mathcal{X}'$ are *equivalent* if there exists a nonsingular linear transformation (resp. matrix) T such that $T^{-1}\mathcal{X}(g)T = \mathcal{X}'(g)$ for all $g \in G$. Given a representation $\mathcal{X}$, we often consider the function $\chi : G \rightarrow \mathbf{C}$ given by $\chi(g) = \text{tr}(\mathcal{X}(g))$. We call χ the *character afforded* by $\mathcal{X}$. Clearly, equivalent representations afford the same character. Let $\mathbf{CG}$ be the complex vector space whose basis is the set G . We convert $\mathbf{CG}$ into a complex algebra by extending the group multiplication linearly. That is,

$$\left(\sum_{g \in G} \alpha(g)g \right) \left(\sum_{h \in G} \beta(h)h \right) = \sum_{g, h} \alpha(g)\beta(h)gh = \sum_x \sum_g \alpha(x) \beta(xg^{-1})x.$$

The resulting algebra is called the *group algebra*. Evidently, $\mathbf{CG}$ is just the algebra of complex-valued functions on G with the convolution product. It is customary, however, to use the formal sums as above. The reason for introducing the group algebra is the useful observation that to every finite-dimensional (right) $\mathbf{CG}$ -module V there is naturally associated a linear representation $\mathcal{X}$ of G . Namely, $\mathcal{X}(g)$ is just right multiplication by g for any $g \in G$. We say that $\mathcal{X}$ is *afforded* by V . Conversely, any linear representation $\mathcal{X} : G \rightarrow \mathbf{GL}(V)$ can be extended linearly to a homomorphism $\mathcal{X} : \mathbf{CG} \rightarrow \text{End}(V)$, thereby converting V into a $\mathbf{CG}$ -module. Clearly, linear representations and $\mathbf{CG}$ -modules are naturally equivalent gadgets.

A representation $\mathcal{X} : G \rightarrow \mathbf{GL}(V)$ is called *reducible* if the corresponding $\mathbf{CG}$ -module is reducible, otherwise it is *irreducible*. An *irreducible character* is the character of an irreducible representation. A particularly important char-

acter is the so-called *regular character* ρ_G afforded by (right multiplication on) the group algebra itself.

$$(2.1) \quad \rho_G(g) = \begin{cases} 0 & \text{if } g \neq 1 \\ |G| & \text{if } g = 1 \end{cases}. \text{ In particular, } CG \text{ is semisimple.}$$

PROOF. Calculating with respect to the basis of group elements, it is immediate that the trace of g_R is zero unless $g = 1$. In particular, the matrix of the trace form is $|G|P$ where P is the matrix of the permutation $g \rightarrow g^{-1}$. $\square$

As a consequence of (2.1) and (1.6), every CG -module is a direct sum of irreducibles. Hence, every character is a sum of irreducible characters.

By (1.3), we can choose notation for the remainder of this section as follows:

- $\{e_1, \dots, e_s\}$ are the primitive central idempotents of CG ,
- $B_i = e_i CG$ ($1 \leq i \leq s$) are the minimal 2-sided ideals of CG ,
- $I_i \subseteq B_i$ is a minimal right ideal affording the irreducible character χ_i ($1 \leq i \leq s$).

Moreover, since the complex numbers are algebraically closed, (1.5) implies that $B_i \cong M(n_i, \mathbb{C})$ where $n_i = \chi_i(1)$. The integer $\chi(1)$ is usually called the *degree* of χ .

(2.2) *Let ρ be the trace of the regular representation of $M(n, \mathbb{C})$ acting on itself by right multiplication. Then $\rho(X) = n \cdot \text{tr}(X)$ for any matrix X .*

PROOF. Calculating with respect to the basis of matrix units E_{ij} , let $X = \sum_{i,j} x_{ij} E_{ij}$. Then we have $E_{ij}X = \sum_k x_{jk} E_{ik}$, whence

$$\rho(X) = \sum_{i,j} x_{jj} = n \cdot \text{tr}(X). \quad \square$$

Now define $\rho_i(\alpha) = \rho_G(e_i \alpha)$ for any $\alpha \in CG$, and notice that $\rho_i(g)$ is just the trace of right multiplication by $e_i g$ on B_i . If we choose a basis for I_i and let $\mathbf{X}_i : B_i \rightarrow M(n_i, \mathbb{C})$ be the homomorphism induced by right multiplication, then $\chi_i(g) = \text{tr}(\mathbf{X}_i(g))$ for any $g \in G$. But $\mathbf{X}_i$ is an isomorphism by (1.5), hence (2.2) implies

$$(2.3) \quad \rho_G(e_i g) = \chi_i(1) \chi_i(g) \quad \text{for all } g \in G \text{ and } 1 \leq i \leq s. \quad \square$$

We can use this result to express the e_i as linear combinations of group elements.

(2.4) *For $i = 1, 2, \dots, s$ we have:*

$$e_i = \frac{\chi_i(1)}{|G|} \sum_{g \in G} \chi_i(g^{-1}) g.$$

PROOF. There are uniquely defined complex numbers $\varepsilon_i(x)$ such that $e_i = \sum_{x \in G} \varepsilon_i(x) x$. Multiplying both sides by g we get

$$e_i g = \sum_{x \in G} \varepsilon_i(x) x g = \sum_{x \in G} \varepsilon_i(x g^{-1}) x.$$

We now apply ρ_G to both sides of this and use (2.1) and (2.3) to obtain

$$\chi_i(1)\chi_i(g) = \rho_G(e_i g) = \sum_{x \in G} \varepsilon_i(x g^{-1}) \rho_G(x) = |G| \varepsilon_i(g^{-1})$$

and the result follows. $\square$

As an easy corollary of (2.4) we obtain the so-called *first orthogonality relation*.

(2.5) **FIRST ORTHOGONALITY RELATION.** *Let χ_i and χ_j be irreducible characters of G . Then for any $x \in G$,*

$$\frac{1}{|G|} \sum_{g \in G} \chi_i(g^{-1}) \chi_j(xg) = \delta_{ij} \frac{\chi_i(x)}{\chi_i(1)}.$$

PROOF. Use (2.4) to substitute into the equation $e_i e_j = \delta_{ij} e_i$:

$$\frac{\chi_i(1)\chi_j(1)}{|G|^2} \sum_{g, h} \chi_i(g^{-1}) \chi_j(h^{-1}) gh = \delta_{ij} \frac{\chi_i(1)}{|G|} \sum_x \chi_i(x^{-1})x.$$

The result follows by equating coefficients of x^{-1} . $\square$

The above result is most often applied in the special case $x = 1$.

Now let $\mathbf{Z}(\mathbf{C}G)$ be the center of the group algebra. Since the group elements are a basis for $\mathbf{C}G$, a necessary and sufficient condition for $\alpha \in \mathbf{C}G$ to lie in $\mathbf{Z}(\mathbf{C}G)$ is that $g^{-1}\alpha g = \alpha$ for all $g \in G$. If $\alpha = \sum_x a(x)x$, then $g^{-1}\alpha g = \sum_x a(gxg^{-1})x$, so $\alpha \in \mathbf{Z}(\mathbf{C}G)$ iff a is constant on conjugacy classes of G . If we therefore let $\hat{x}$ denote the sum of all G -conjugates of x in G , then the distinct sums $\hat{x}$ as x ranges over G form a basis for $\mathbf{Z}(\mathbf{C}G)$. Indeed, the same argument shows more: the class sums are a basis for the *integral* group ring $\mathbf{Z}G \subseteq \mathbf{C}G$. But the primitive central idempotents also form a basis for $\mathbf{Z}(\mathbf{C}G)$ since

$$\mathbf{Z}(\mathbf{C}G) = \bigoplus_{i=1}^s \mathbf{Z}(B_i)$$

and $\mathbf{Z}(M(n, \mathbf{C}))$ is just the scalar matrices. Since the χ_i are in 1-1 correspondence with the e_i , we have proved

(2.6) *The distinct conjugacy class sums and the primitive central idempotents are both bases for $\mathbf{Z}(\mathbf{C}G)$. In particular, the number of irreducible characters is equal to the number of conjugacy classes.* $\square$

We next take a closer look at the character values. Suppose that χ is afforded by a matrix representation $\mathcal{X}$, and let σ be an automorphism of the complex numbers. If we denote by $\mathcal{X}^\sigma(g)$ the result of applying σ to the matrix entries of $\mathcal{X}(g)$, it is clear that $\mathcal{X}^\sigma$ is also a representation, which is irreducible if $\mathcal{X}$ is. Thus, the function $\chi^\sigma(g) = \sigma(\chi(g))$ is another (not necessarily distinct) character which is irreducible if χ is.

(2.7) *Let $\mathcal{X}$ be a representation of G of degree n affording the character χ and let $g \in G$. Then:*

- (i) $\chi(g)$ is a sum of $n|G|$ th roots of unity. In particular, $|\chi(g)| \leq \chi(1)$ with equality iff $\mathcal{X}(g)$ is a scalar matrix, and $\chi(g) = \chi(1)$ iff $\mathcal{X}(g) = I$.

- (ii) *For any automorphism σ of the complex numbers, there exists an integer i relatively prime to $|G|$ and depending only on σ such that $\chi^\sigma(g) = \chi(g^i)$ for all $g \in G$. If σ is complex conjugation, we may take $i = -1$.*

PROOF. Choose $g \in G$. Then $g^e = 1$ for some divisor e of $|G|$. It follows that $\mathcal{X}(g)$ satisfies the polynomial $\lambda^{|G|} - 1$. In particular, the minimum polynomial of $\mathcal{X}(g)$ has distinct roots which are certain $|G|$ th roots of unity. Thus, after replacing $\mathcal{X}$ by a similar representation if necessary, we may assume that $\mathcal{X}(g)$ is a diagonal matrix with $|G|$ th roots of unity on the diagonal. The triangle inequality implies that the sum of n roots of unity has absolute value less than n with equality iff they are all equal. Since $\chi(1) = n$ we have $|\chi(g)| \leq \chi(1)$ with equality iff $\mathcal{X}(g)$ is a scalar matrix. Clearly then, if $\chi(g) = \chi(1)$, that scalar matrix must be the identity.

Let ξ be a primitive $|G|$ th root of unity. Then $\sigma(\xi) = \xi^i$ for some integer i relatively prime to $|G|$ and depending only on σ . Since $\mathcal{X}(g)$ is diagonal with powers of ξ on the diagonal, it is clear that $\mathcal{X}^\sigma(g) = \mathcal{X}(g^i)$ and therefore $\chi^\sigma(g) = \chi(g^i)$ as required. $\square$

From the above, we see that the set of $g \in G$ with $\chi(g) = \chi(1)$ is a normal subgroup of G , which we denote by $\ker(\chi)$.

By a *class function* we mean a complex-valued function on G which is constant on conjugacy classes. We define a positive definite Hermitian inner product on the space of class functions as follows:

$$(\eta, \varphi) = \frac{1}{|G|} \sum_{g \in G} \overline{\eta(g)} \varphi(g).$$

(2.8) **SECOND ORTHOGONALITY RELATION.** *The irreducible characters form an orthonormal basis for the space of class functions. Let x_j be a representative of the j th conjugacy class of G and let $\mathbf{C}_G(x_j)$ denote the centralizer of x_j in G . Then*

$$\sum_{i=1}^s \chi_i(x_j) \overline{\chi_i(x_k)} = \delta_{jk} |\mathbf{C}_G(x_j)|.$$

PROOF. From (2.7) we have $\chi(g^{-1}) = \overline{\chi(g)}$, then (2.5) (with $x = 1$) says that the irreducible characters are an orthonormal set. From (2.6) it then follows that the irreducible characters are a basis. Let $\mathbf{X}$ be the $s \times s$ matrix whose (i, j) entry is $\chi_i(x_j)$. $\mathbf{X}$ is called the *character table* of G . Since the χ_i are class functions, the first orthogonality relation can be written

$$\frac{1}{|G|} \sum_{k=1}^s |\hat{x}_k| \overline{\chi_i(x_k)} \chi_j(x_k) = \delta_{ij}.$$

If D is the diagonal matrix with (k, k) entry $|\hat{x}_k|$, the above equation can be written in matrix form:

$$\mathbf{X}^* D \mathbf{X} = |G| I$$

where $*$ denotes conjugate transpose. Then

$$\mathbf{X} \mathbf{X}^* D = \mathbf{X} (\mathbf{X}^* D \mathbf{X}) \mathbf{X}^{-1} = \mathbf{X} (|G| I) \mathbf{X}^{-1} = |G| I,$$

whence

$$\mathbf{X}\mathbf{X}^* = |G|D^{-1}.$$

Since $|\hat{x}_k| = |G : \mathbf{C}_G(x_k)|$, the proof is complete. $\square$

We let $\text{Irr}(G)$ denote the set of irreducible characters of G . Since $\text{Irr}(G)$ is an orthonormal basis, any class function ϕ has a “Fourier expansion”:

$$\phi = \sum_{\chi \in \text{Irr}(G)} (\chi, \phi) \chi.$$

It follows that

(2.9) *A class function ϕ is a character iff (χ, ϕ) is a nonnegative integer for all $\chi \in \text{Irr}(G)$. A character ϕ is irreducible iff $(\phi, \phi) = 1$. $\square$*

We next record an immediate consequence of (1.9).

(2.10) *Suppose that K and H are finite groups. Then every irreducible character of $K \times H$ is of the form $\chi_{\psi\phi}(k, h) = \psi(k)\phi(h)$ where $\psi \in \text{Irr}(K)$ and $\phi \in \text{Irr}(H)$. $\square$*

A class function ϕ for which (χ, ϕ) is an integer (possibly negative) is called a *generalized* (or *virtual*) *character*. An important observation here is that the set of generalized characters actually forms a ring under pointwise multiplication. This is immediate from

(2.11) *The pointwise product of characters is a character.*

PROOF. Embed G into $G \times G$ on the diagonal. Then the restriction of the irreducible character $\chi_{\phi\psi}$ of $G \times G$ (see (2.10)) is a character of G . $\square$

We conclude this section by looking at several examples. First of all, any 1-dimensional $\mathbf{C}G$ module is clearly irreducible, and every group has at least one such module, namely the trivial module where all group elements fix all vectors. The character afforded by this module is called the *principal character*, and is denoted 1_G . It has the value 1 at all group elements.

More generally, characters of degree 1 are called *linear characters*. They are just homomorphisms $G \rightarrow \mathbf{C}^\times$. Suppose that G is abelian. Then $\mathbf{C}G$ is a commutative algebra, so by (1.2) and (1.3), all the irreducible characters are linear. The converse of this statement is also true.

An important class of examples of characters are the *permutation characters*. If G acts on a set Ω then the vector space $\mathbf{C}\Omega$ with basis Ω affords a representation of G . The resulting character gives the number of fixed points of each group element on Ω . Notice that $\mathbf{C}\Omega$ always has a 1-dimensional trivial submodule spanned by the sum of the basis vectors, so there is a codimension 1 complement. For example, S_n has a character of degree $n - 1$ (which, as we shall see, is irreducible).

CHAPTER 3

Divisibility

In this section, we obtain some nontrivial results by considering some divisibility properties of character values. First, we need to recall a few algebraic facts. A complex number γ is called an *algebraic integer* if it is the root of a monic polynomial with integer coefficients.

(3.1) *The set of algebraic integers is a subring of $\mathbb{C}$ whose intersection with the rational numbers is the integers.*

PROOF. See any basic algebra text (e.g., [7]).

Now let $g \in G$ and let $\hat{g}$ be the conjugacy class sum defined in §2. Then by (2.6) there are uniquely defined complex numbers $\omega_i(\hat{g})$ such that $\hat{g} = \sum_{i=1}^s \omega_i(\hat{g})e_i$. We extend the ω_i linearly to complex-valued functions on $\mathbf{Z}(\mathbf{C}G)$. By $|\hat{g}|$ we mean the number of terms in the sum, i.e., the number of conjugates of g .

(3.2) *The functions $\omega_i: \mathbf{Z}(\mathbf{C}G) \rightarrow \mathbb{C}$ are algebra homomorphisms whose values are algebraic integers. Moreover,*

$$\omega_i(\hat{g}) = \frac{|\hat{g}|\chi_i(g)}{\chi_i(1)}.$$

PROOF. The fact that the ω_i are algebra homomorphisms follows immediately from the fact that the e_i are orthogonal idempotents. To see that $\omega_i(\hat{g})$ is an algebraic integer, we use an observation made earlier that the $\hat{g}$ are in fact a basis for the center of the integral group ring $\mathbf{Z}G$. Hence, there are rational integers a_{ijk} such that

$$\hat{x}_i \hat{x}_j = \sum_k a_{ijk} \hat{x}_k.$$

Fixing r and applying ω_r to this equation we obtain

$$(3.3) \quad \omega_r(\hat{x}_i) \omega_r(\hat{x}_j) = \sum_k a_{ijk} \omega_r(\hat{x}_k).$$

Let A_i be the $s \times s$ integral matrix with (j, k) entry a_{ijk} and let w_r be the vector whose j th entry is $\omega_r(\hat{x}_j)$. Then (3.3) becomes

$$A_i w_r = \omega_r(\hat{x}_i) w_r.$$

In particular, $\omega_r(\hat{x}_i)$ is a root of the characteristic polynomial of A_i which is a monic polynomial with integer coefficients, and thus $\omega_r(\hat{x}_i)$ is an algebraic integer.

To obtain the desired formula for the ω_i in terms of χ_i we use the functions ρ_i of (2.3). Since characters are traces, they are constant on conjugacy classes, so that

$$\rho_i(\hat{g}) = |\hat{g}| \chi_i(1) \chi_i(g)$$

by (2.3). On the other hand $\hat{g}e_i = \omega_i(g)e_i$, whence (2.3) yields

$$\rho_i(\hat{g}) = \omega_i(g) \rho_i(e_i) = \omega_i(g) \rho_i(1) = \omega_i(g) \chi_i(1)^2.$$

Equating these two expressions completes the proof. $\square$

(3.4) *The degree of an irreducible character divides the order of the group.*

PROOF. Fix $\chi \in \text{Irr}(G)$. From the orthogonality relations, we have

$$|G| = \sum_{g \in G} \chi(g) \overline{\chi(g)}.$$

Choosing conjugacy class representatives $\{x_1, \dots, x_s\}$ we can rewrite this as

$$|G| = \sum_{i=1}^s |\hat{x}_i| \chi(x_i) \overline{\chi(x_i)}.$$

Dividing by $\chi(1)$ and using (3.2) we get

$$\frac{|G|}{\chi(1)} = \sum_{i=1}^s \omega(\hat{x}_i) \overline{\chi(x_i)}.$$

Since roots of unity are obviously algebraic integers, the right-hand side is an algebraic integer by (3.2), (2.7), and (3.1). Hence, the left-hand side is an integer by (3.1). $\square$

(3.5) *Suppose $\chi \in \text{Irr}(G)$ and $x \in G$ such that $\gcd(\chi(1), |\hat{x}|) = 1$. Then either $\chi(x) = 0$ or $x \in \mathbf{Z}(G/\ker(\chi))$.*

PROOF. Choose integers a and b such that $a\chi(1) + b|\hat{x}| = 1$. Then

$$\frac{\chi(x)}{\chi(1)} = \frac{\chi(x)}{\chi(1)} (a\chi(1) + b|\hat{x}|) = a\chi(x) + b \frac{|\hat{x}| \chi(x)}{\chi(1)},$$

whence $\chi(x)/\chi(1)$ is an algebraic integer by (3.2). If k is any integer relatively prime to $|G|$, then $C_G(x^k) = C_G(x)$, so in particular $|\hat{x}| = |\widehat{x^k}|$. Thus, the above argument may be repeated with x^k in place of x . By (2.7) we conclude that all Galois conjugates of $\frac{\chi(x)}{\chi(1)}$ are also algebraic integers, and each has absolute value at most one. Thus, the Galois norm of $\frac{\chi(x)}{\chi(1)}$ is a rational integer of absolute value at most one, so if $\chi(x) \neq 0$ we get $|\chi(x)| = \chi(1)$. By (2.7), a representation $\mathcal{R}$ affording χ embeds $G/\ker(\chi)$ into $M_n(\mathbf{C})$ in such a way that $\mathcal{R}(x)$ is a central element. $\square$

3.6 (Burnside). (i) Suppose $|\hat{x}| = p^r$ for some nonidentity element $x \in G$ and some prime p . Then G is not simple.

(ii) Every group of order $p^a q^b$ (p and q primes) is solvable.

PROOF. (i) Let ρ_G be the regular character and 1_G the principal character. Then

$$\rho_G(x) = 0 = 1 + \sum_{\chi \neq 1} \chi(1)\chi(x).$$

It follows that there is a nonprincipal character χ such that $\chi(x) \neq 0$ and $p \nmid \chi(1)$; otherwise the above equation would imply that $1/p$ is an algebraic integer. Now (3.5) implies that $x \in \mathbf{Z}(G/\ker(\chi))$, so G cannot be simple. $\square$

(ii) If $|G| = p^a q^b$, let Q be a Sylow q -subgroup of G and choose a nonidentity element $x \in \mathbf{Z}(Q)$. Then $Q \subseteq \mathbf{C}_G(x)$ which implies that $|\hat{x}| = p^r$ for some integer $r \leq a$. By the first paragraph, either G is of prime order or G has a proper normal subgroup. Hence G is solvable by an obvious induction argument. $\square$

CHAPTER 4

Induced Characters

Let $H \subseteq G$ and let ϕ be a class function on H . Extend ϕ to a function $\dot{\phi}$ on G by defining

$$\dot{\phi}(g) = \begin{cases} \phi(g), & g \in H, \\ 0, & g \notin H. \end{cases}$$

Now define *the induced class function* ϕ^G on G as follows:

$$\phi^G(g) = \frac{1}{|H|} \sum_{x \in G} \dot{\phi}(xgx^{-1}).$$

We denote the restriction of a function ψ on G to H by ψ_H . We will be interested mostly in the case where ϕ is either a character or at worst a *generalized character*, by which we mean an integral linear combination of characters.

(4.1) **FROBENIUS RECIPROCITY.** *Let ϕ be a class function on $H \subseteq G$ and let ψ be a class function on G . Then*

$$(\psi, \phi^G) = (\psi_H, \phi).$$

PROOF. This is a straightforward calculation:

$$(\psi, \phi^G) = \frac{1}{|G||H|} \sum_{g, x} \bar{\psi}(g) \dot{\phi}(xgx^{-1}) = \frac{1}{|G||H|} \sum_x \sum_g \bar{\psi}(x^{-1}gx) \dot{\phi}(g).$$

Since $\bar{\psi}$ is a class function and $\dot{\phi}$ vanishes off H , we get

$$(\psi, \phi^G) = \frac{1}{|G||H|} \sum_x \sum_{h \in H} \bar{\psi}(h) \phi(h) = (\psi_H, \phi). \quad \square$$

The following corollary is immediate from (2.9) and (4.1).

(4.2) *If ϕ is a (generalized) character of a subgroup $H \subseteq G$ then ϕ^G is a (generalized) character of G . $\square$*

It turns out that if V is a CH -module affording ϕ , we can use the standard tensor product construction to extend the ring of operators:

$$V^G = V \otimes_{CH} CG,$$

where we are regarding CG as a left CH -module and a right CG -module. Then V^G affords ϕ^G . Since we do not need this result, we omit the proof.

The induction map is indispensable in analyzing the relationship between characters of G and characters of subgroups $H \subseteq G$. Here are some of its useful properties:

(4.3) Let ϕ be a class function on $H \subseteq G$ and let ψ be a class function on G . Then:

(a) $(\psi_H \phi)^G = \psi \phi^G$.

(b) Let $x_1, \dots, x_t$ be a set of right coset representatives for H in G . Then

$$\phi^G(g) = \sum_{i=1}^t \phi(x_i g x_i^{-1}).$$

(c) If $H \subseteq K \subseteq G$, then $(\phi^K)^G = \phi^G$.

PROOF. From the definition, we have

$$(\psi_H \phi)^G(g) = \frac{1}{|H|} \sum_{x \in G} \psi_H(x g x^{-1}) \phi(x g x^{-1}) = \frac{1}{|H|} \sum_{x \in G} \psi(x g x^{-1}) \phi(x g x^{-1}),$$

but since ψ is constant on G conjugacy classes, assertion (a) follows.

Statement (b) follows immediately from the definition and the fact that ϕ is constant on H conjugacy classes.

The last statement can be proved easily from the definitions, but we note that it is immediate from (4.1), (2.8), and the trivial fact that $(\chi_K)_H = \chi_H$. $\square$

4.4 (Mackey). Let $K, H \subseteq G$, suppose that ψ is a class function on K , and let $x_1, \dots, x_t$ be a set of (K, H) double coset representatives in G . For each i , let $K_i = x_i^{-1} K x_i$ and define class functions $\psi^{(i)}(y) = \psi(x_i y x_i^{-1})$ on K_i . Put $H_i = H \cap K_i$. Then

$$(\psi^G)_H = \sum_{i=1}^t (\psi_{H_i}^{(i)})^H.$$

In particular, if ϕ is a class function on H , then

$$(\phi^G, \psi^G) = \sum_{i=1}^t (\phi_{H_i}, \psi_{H_i}^{(i)}).$$

PROOF. The second conclusion follows from the first by Frobenius reciprocity:

$$(\phi^G, \psi^G) = (\phi, (\psi^G)_H) = \sum_{i=1}^t (\phi, (\psi_{H_i}^{(i)})^H) = \sum_{i=1}^t (\phi_{H_i}, \psi_{H_i}^{(i)}).$$

To prove the first statement, consider the action of G by right multiplication on the right cosets Kx of K . The double coset $Kx_i H$ is the orbit of H containing the point Kx_i , and the stabilizer of this point in H is precisely H_i . Let $\{h_{i1}, \dots, h_{iT_i}\}$ be a set of right coset representatives for H_i in H . Then each point Kx in the H -orbit $Kx_i H$ can be written $Kx_i h_{ij}$ for some j . In particular, the set of products $\{x_i h_{ij} \mid 1 \leq i \leq t, 1 \leq j \leq T_i\}$ is a set of right

coset representatives for K in G . Let $h \in H$; then by (4.3)(b) we have

$$\psi^G(h) = \sum_{i=1}^t \sum_{j=1}^{t_i} \psi(x_i h_{ij} h h_{ij}^{-1} x_i^{-1}) = \sum_{i=1}^t \sum_{j=1}^{t_i} \psi^{(i)}(h_{ij} h h_{ij}^{-1}),$$

where $\psi^{(i)}$ vanishes off K_i . But then (4.3)(b) applied to the subgroup $H_i \subseteq H$ shows that the inner sum is precisely $(\psi_{H_i}^{(i)})^H(h)$. $\square$

We already have enough information about induced characters to prove the following famous theorem of Frobenius, for which no purely group-theoretic proof is known.

4.5 (Frobenius). *Suppose $H \subseteq G$ and $H \cap g^{-1}Hg = 1$ for $g \in G \setminus H$. Then there exists a normal subgroup $N \trianglelefteq G$ with $H \cap N = 1$ and $HN = G$.*

PROOF. Let S be the union of all the conjugates of the set $H \setminus \{1\}$ in G . Since our hypothesis implies in particular that $N_G(H) = H$, it follows that S is the union of $|G : H|$ disjoint sets, each of which has cardinality $|H| - 1$, and therefore $|S| = |G : H|(|H| - 1)$. Let $N = G \setminus S$; then $|N| = |G : H|$. Moreover, it is clear that N is a union of G -conjugacy classes. The problem is to prove that N is a subgroup.

In order to do this, we let $\text{Irr}(H) = \{\phi_0 = 1_H, \phi_1, \dots, \phi_s\}$ and define the generalized characters $\psi_i = \phi_i(1)1_H - \phi_i$ ($1 \leq i \leq s$). Then for $1 \leq i, j \leq s$ we have

$$(4.6) \quad (\psi_i, \psi_j) = \phi_i(1)\phi_j(1) + \delta_{ij}, \quad (1_H, \psi_i) = \phi_i(1), \quad \psi_i(1) = 0.$$

Now consider the generalized characters ψ_i^G of G . It is immediate from the definition of induced characters that $\psi_i^G(1) = 0$, whence the Mackey theorem (4.4) with $H = K$ and our hypothesis imply that

$$(4.7) \quad (\psi_i^G, \psi_j^G) = (\psi_i, \psi_j) \quad (1 \leq i, j \leq s).$$

By Frobenius reciprocity, we easily get $(1_G, \psi_i^G) = (1_H, \psi_i) = \phi_i(1)$, which means that we can write $\psi_i^G = \phi_i(1)1_G - \chi_i$ where χ_i is some generalized character with $(\chi_i, 1_G) = 0$. But now, (4.6) and (4.7) yield $(\chi_i, \chi_j) = \delta_{ij}$. In particular, χ_i must be (up to sign) an irreducible character. Since $\psi_i^G(1) = 0$, we must have $\chi_i(1) = \phi_i(1) > 0$ and hence $\{\chi_1, \dots, \chi_s\}$ is a set of distinct irreducible characters of G . Let N_0 be the intersection of their kernels. We will show that $N_0 = N$.

Applying the Mackey theorem again we see that

$$\phi_i(1)1_H - \chi_{iH} = (\psi_i^G)_H = \psi_i = \phi_i(1)1_H - \phi_i$$

and therefore that $\chi_{iH} = \phi_i$ for all i . Since the intersection of the kernels of the nonprincipal irreducible characters of any group is the identity by (2.8), we have $N_0 \cap H = 1$.

On the other hand, if $x \in N \setminus \{1\}$ then x is not G -conjugate to any element of H , whence $\psi_i^G(x) = 0$ for each i by definition of induced characters. But this implies that $x \in \ker(\chi_i)$ for all i and thus $N \subseteq N_0$. Since $|G| = |H||N| \geq |H||N_0|$ and $H \cap N_0 = 1$, it follows that $G = HN$ and $N = N_0$. $\square$

A very important class of examples of induced characters is provided by the permutation characters. The details are as follows:

(4.8) Suppose that G acts on a set Ω with permutation character θ . Let $\Omega_1, \dots, \Omega_r$ be the G -orbits on Ω , and let H_i be the stabilizer of a point in Ω_i ($1 \leq i \leq r$). Then $\theta = \sum_{i=1}^r (1_{H_i})^G$. In particular, $(\theta, 1_G) = r$. If $r = 1$ then (θ, θ) is the number of double cosets of H_1 in G .

PROOF. Since θ is the sum of the transitive permutation characters θ_i of G acting on Ω_i , it suffices to consider the special case $r = 1$, $H = H_1$. So we may assume that Ω is the set of right cosets of H with G acting by right multiplication. Then we see from (4.3)(b) that $1_H^G(g)$ is the number of right coset representatives x_j of H in G for which $x_j g x_j^{-1}$ lies in H . But this is just the number of cosets Hx_j with $Hx_j g = Hx_j$.

We have shown that $\theta_i = (1_{H_i})^G$. By Frobenius reciprocity, $(\theta_i, 1_G) = 1$ so $(\theta, 1_G) = r$. If $r = 1$ and H is a point stabilizer, then (4.4) shows that (θ, θ) is the number of (H, H) double cosets. $\square$

The above result says that the number of orbits of a group acting on a set is equal to the average number of fixed points. This observation, originally due to Burnside, is useful in certain enumeration problems.

Another consequence of (4.8) is the case that θ is the character of a doubly transitive permutation representation. In that case, it is easy to see that there are exactly two double cosets of a point stabilizer, so we get $(\theta, \theta) = 2$. This implies that $\theta = 1_G + \chi$ where χ is irreducible.

As mentioned before, we will not need to use induced modules very much here, relying for the most part on the simpler induced characters. However, the following is one special case of interest:

(4.9) Suppose $H \subseteq G$ and λ is a linear character of H with corresponding central idempotent e_λ . Then the principal right ideal $e_\lambda \mathbf{C}G$ of $\mathbf{C}G$ affords the induced character λ^G .

PROOF. Let $\{x_1, x_2, \dots, x_t\}$ be a set of right coset representatives for H in G . Then

$$e_\lambda \mathbf{C}G = \mathbf{C}e_\lambda x_1 + \dots + \mathbf{C}e_\lambda x_t$$

because $e_\lambda h = \lambda(h)e_\lambda$ for all $h \in H$. Moreover, since

$$e_\lambda x_i = \frac{1}{|H|} \sum_{h \in H} \lambda(h^{-1}) h x_i,$$

the vectors $B = \{e_\lambda x_i \mid 1 \leq i \leq t\}$ are a linear basis for the ideal $e_\lambda \mathbf{C}G$ since the sets Hx_i are disjoint. Given any $g \in G$, there is a permutation $i \rightarrow i'$ of $\{1, 2, \dots, t\}$ and elements $h_i(g) \in H$ such that for each i , $x_i g = h_i(g) x_{i'}$. This means that the matrix of right multiplication by g is monomial with respect to the basis B , the unique nonzero element in row i being $\lambda(h_i(g))$ in column i' . We get a nonzero contribution to the trace precisely when $i = i'$, i.e., when $x_i g x_i^{-1} \in H$, and the result now follows from (4.3)(b). $\square$

CHAPTER 5

Further Results

In this section, we obtain a number of important and inter-related results in character theory, including Clifford's theorem on characters of normal subgroups, the fact that all irreducible representations of p -groups are monomial, Brauer's characterization of characters, and Brauer's theorem on blocks of defect zero. None of these results is needed elsewhere in these notes.

We begin with Clifford's theorem. Suppose that $H \trianglelefteq G$ and $\theta \in \text{Irr}(H)$ is afforded by a representation Θ . Then for any $g \in G$, composing the map $h \mapsto ghg^{-1}$ with Θ yields another representation affording the character $\theta^g(h) := \theta(ghg^{-1})$. In this way, G acts on $\text{Irr}(H)$ with H acting trivially. For each $\theta \in \text{Irr}(H)$ we put $G_\theta = \{g \in G \mid \theta^g = \theta\}$; then $H \subseteq G_\theta \subseteq G$.

5.1 (Clifford). *Suppose that $H \trianglelefteq G$, $\chi \in \text{Irr}(G)$, and θ is an irreducible constituent of χ_H . Then there is a unique irreducible character ψ of G_θ for which $(\psi^G, \chi) \neq 0 \neq (\psi_H, \theta)$. Moreover, $\psi^G = \chi$, $\psi_H = e\theta$ for some integer e , and if X is a set of coset representatives for G_θ in G then $\chi_H = e \sum_{x \in X} \theta^x$.*

PROOF. Let $I = G_\theta$. Applying (4.4) with $K = H$ yields

$$(\theta^G)_H = \sum_{y \in G/H} \theta^y = |I : H| \sum_{x \in X} \theta^x$$

because (H, H) double cosets are just H -cosets when H is normal. Since χ is a constituent of θ^G by reciprocity, χ_H is a constituent of $(\theta^G)_H$ and we can thus see that all irreducible constituents of χ_H are G -conjugate to θ . On the other hand, χ_H is G -invariant, so $(\chi, \theta) = (\chi, \theta^g)$ for all $g \in G$. Hence

$$\chi_H = e \sum_{x \in X} \theta^x$$

for some integer e . In particular, $\chi(1) = e|G : I|\theta(1)$.

We can write $\chi_I = \psi + \phi$ where every irreducible constituent ξ of ψ satisfies $(\xi_H, \theta) \neq 0$ while $(\phi_H, \theta) = 0$. Then $\psi_H = e\theta$ so that $\psi(1) = e\theta(1)$, and $(\chi_I, \psi) = (\psi, \psi)$. By reciprocity it follows that $\psi^G = (\psi, \psi)\chi + \vartheta$ for some character ϑ of G with $(\chi, \vartheta) = 0$, whence

$$|G : I|e\theta(1) = |G : I|\psi(1) = \psi^G(1) \geq (\psi, \psi)\chi(1) = (\psi, \psi)e|G : I|\theta(1).$$

We conclude that $(\psi, \psi) = 1$ and ψ is therefore the unique irreducible character of I satisfying $(\chi_I, \psi) \neq 0 \neq (\psi_H, \theta)$. Moreover, the inequality is an equality which means that $\vartheta = 0$. $\square$

(5.2) Suppose that G has a normal abelian subgroup A such that G/A is a p -group for some prime p . Then for each $\chi \in \text{Irr}(G)$ there is a subgroup H of G and a linear character λ of H with $\lambda^G = \chi$.

PROOF. We may assume without loss of generality that G is a minimal counterexample and that A is a maximal normal abelian subgroup of G . We first argue that $A = C_G(A)$, for if not then $C_G(A)/A$ is a proper normal subgroup of the p -group G/A and therefore meets the center of G/A nontrivially. This implies that there is a normal subgroup Z of G with $A \subseteq Z \subseteq C_G(A)$ and $|Z/A| = p$. In particular, A is central in Z and Z/A is cyclic, whence Z is abelian contrary to the maximality of A . We conclude that

$$(5.3) \quad A = C_G(A).$$

Since every subgroup of G satisfies the hypothesis, it suffices by induction to show that every nonlinear character χ of G is induced from a proper subgroup of G . Hence, by (5.1) with $H = A$, we may assume that χ is nonlinear and

$$(5.4) \quad \chi_A = e\theta$$

for some irreducible (and hence linear) character θ of A .

Let $N = \ker(\chi)$. Then, by a slight abuse of notation, χ is an irreducible character of G/N and if $\chi = \lambda^{G/N}$ for some linear character λ of a subgroup H/N , it is an easy exercise to see that $\chi = \lambda^G$. Since G/N satisfies the hypothesis, we may assume by induction that $N = 1$.

Now let $\mathcal{X}$ be a representation affording χ . Then $\mathcal{X}$ embeds G into $\text{GL}(V)$ and (5.4) shows that $\mathcal{X}(a)$ is a scalar for all $a \in A$. By (5.3), $A = G$ and hence χ is linear. $\square$

Our next objective in this section is to prove an important result due to Brauer which gives a necessary and sufficient condition for a class function on G to be a generalized character. First, we need some notation and definitions.

Let $\text{Ch}(G)$ be the ring of generalized characters of G . Let $\mathcal{H}$ be a family of subgroups of G with the property that if $H, K \in \mathcal{H}$ and $g \in G$, then $H \cap K^g \in \mathcal{H}$, and let $\mathcal{B}(G; \mathcal{H})$ be the set of permutation characters $\{1_H^G \mid H \in \mathcal{H}\}$.

$$(5.5) \quad \mathcal{B}(G; \mathcal{H}) \text{ is a subring of } \text{Ch}(G).$$

PROOF. Let $H, K \in \mathcal{H}$. By (4.3)(a),(c), and (4.4),

$$(1_H^G)(1_K^G) = ((1_H)(1_K)_H)^G = ((1_K)_H)^G = \left(\sum_{i=1}^t 1_{H_i}^H \right)^G = \sum_{i=1}^t 1_{H_i}^G$$

where $H_i = H \cap x_i^{-1}Kx_i$ as in (4.4). $\square$

The ring $\mathcal{B}(G; \mathcal{H})$ is often called the *Burnside ring* of G relative to $\mathcal{H}$.

(5.6) Let R be a ring of $\mathbf{Z}$ -valued functions on a finite set G with pointwise operations. Suppose that for each prime p and each $g \in G$, there exists a function $f_{g,p} \in R$ with $f_{g,p}(g) \not\equiv 0 \pmod{p}$. Then $1 \in R$.

PROOF. For $g \in G$, let $I_g = \{f(g) : f \in R\} \subseteq \mathbf{Z}$. I_g is an additive subgroup and therefore an ideal of $\mathbf{Z}$. Our hypothesis thus guarantees that $I_g = R$, whence there exists a function $f_g \in R$ with $f_g(g) = 1$. It follows that

$\prod_{g \in G} (1 - f_g) = 0$. By expanding out this product, we obtain 1 as a sum of elements of R . $\square$

We will call a group H *quasi-elementary* if, for some prime p , H is a semi-direct product PC where C is a normal cyclic subgroup of order prime to p and P is a p -group. It is clear that any subgroup of a quasi-elementary group is itself quasi-elementary, and in particular that the Burnside ring $\mathcal{B}(G; \mathcal{Q})$ is defined for the set $\mathcal{Q}$ of quasi-elementary subgroups of G .

$$(5.7) \quad 1 \in \mathcal{B}(G; \mathcal{Q}).$$

PROOF. It suffices to show that $\mathcal{B}(G; \mathcal{Q})$ satisfies the hypotheses of (5.6). Thus, given a prime p and an element $g \in G$, write the order of g as $p^a n$ where $p \nmid n$ and let $C = \langle g^{p^a} \rangle$. Then $|C| = n$. Let P be a Sylow p -subgroup of $N = N_G(C)$ containing g , and let $H = PC$. Then $H \in \mathcal{Q}$ and we claim that $1_H^G(g) \not\equiv 0 \pmod{p}$.

Namely, choose coset representatives $\{x_1, \dots, x_l\}$ for H in G . Then by (4.3)(b), $1_H^G(g)$ equals the number of indices i for which $x_i g x_i^{-1} \in H$. Now if $x g x^{-1} \in H$, then $x C x^{-1} \subseteq H$ but since H/C is a p -group, C contains all subgroups of H of order prime to p . It follows that $x C x^{-1} = C$, and thus $x g x^{-1} \in H$ implies that $x \in N$. We conclude that $1_H^G(g) = 1_H^N(g)$ is the number of cosets of H in N fixed by g .

Since $C \trianglelefteq N$ and $C \subseteq H$, C fixes all the cosets of H in N and thus the action of g on the cosets of H in N has order dividing p^a . In particular, every nontrivial orbit of g has length divisible by p . On the other hand, the number of cosets of H in N is prime to p because H contains a Sylow p -subgroup of N , and thus the number of fixed points of g must be prime to p as required. $\square$

We now turn to the proof of Brauer's characterization of characters. We say that a subgroup $H \subseteq G$ is *elementary* if, for some prime p , $H = P \times C$ where C is cyclic of order prime to p and P is a p -group. In particular, elementary subgroups are quasi-elementary.

(5.8) **BRAUER'S CHARACTERIZATION OF CHARACTERS.** *Let ϕ be a class function on G . Then the following statements are equivalent:*

- (a) *There exist elementary subgroups H_i , linear characters λ_i of H_i , and integers a_i ($1 \leq i \leq n$) such that $\phi = \sum_{i=1}^n a_i \lambda_i^G$.*
- (b) *ϕ is a generalized character of G .*
- (c) *ϕ_H is a generalized character of H for every elementary subgroup H of G .*

PROOF. Let $\mathcal{E}$ be the set of all elementary subgroups of G . Let $\mathcal{R}$ be the ring of all class functions ϕ on G such that $\phi_H \in \text{Ch}(H)$ for all $H \in \mathcal{E}$, and let $\mathcal{S}$ be the subgroup of $\text{Ch}(G)$ spanned over $\mathbb{Z}$ by characters of the form λ^G where λ is a linear character of some $H \in \mathcal{E}$. Then it is clear that $\mathcal{S} \subseteq \text{Ch}(G) \subseteq \mathcal{R}$, and the theorem is equivalent to the statement $\mathcal{S} = \mathcal{R}$.

Let $\phi \in \mathcal{S}$ and $\psi \in \mathcal{R}$, with $\phi = \sum_{i=1}^n a_i \lambda_i^G$ where λ_i is a linear character of the elementary subgroup H_i . Then $\psi\phi = \sum_{i=1}^n a_i (\psi_{H_i} \lambda_i)^G$ by (4.3)(a). Since

$\psi_{H_i} \in \text{Ch}(H_i)$, there exist integers b_{ij} such that

$$(5.9) \quad \psi\phi = \sum_{i,j} a_i b_{ij} \xi_{ij}^G$$

where $\xi_{ij} \in \text{Irr}(H_i)$. By (5.2), ξ_{ij} is induced from a linear character of some subgroup of H_i , but since subgroups of elementary groups are again elementary, (4.3)(c) applied to (5.9) shows that $\psi\phi \in \mathcal{J}$. This means that $\mathcal{J}$ is an ideal of $\mathcal{R}$, so to complete the proof it suffices to show that $1_G \in \mathcal{J}$.

By (5.7) and (4.3)(c) it suffices to assume that $G = PC$ where C is a normal cyclic subgroup of order prime to p and P is a p -group for some prime p , and then show that $1_G \in \mathcal{J}$. Let $N = \mathbf{N}_G(P)$. Then $N = P \times (N \cap C)$ is elementary. If $N = G$ then G is elementary and there is nothing to prove. So we may as well assume that $N < G$. Let

$$(5.10) \quad 1_N^G = a_0 1_G + \sum_{i>0} a_i \chi_i$$

where the χ_i are nonprincipal irreducible characters and the a_i are positive integers. Notice that $a_0 = (1_N^G, 1_G) = (1_N, 1_N) = 1$.

We next argue that $\chi_i(1) > 1$ for all $i > 0$. Namely, $(\chi_{iN}, 1_N) \neq 0$ by reciprocity, so if χ_i were linear for some i we would have $\chi_{iN} = 1_N$ and N would be contained in the proper normal subgroup $H = \ker(\chi_i)$. But this is impossible by the so-called ‘‘Frattini argument’’:

Let $g \in G$. Then P and P^g are both Sylow p -subgroups of H , whence $P^{g^h} = P$ for some $h \in H$. But then $gh \in N \subseteq H$ and therefore $g \in H$ for all $g \in G$ which is a contradiction.

We can now complete the proof by induction, because by (5.2) each χ_i is induced from a linear character λ_i of a proper subgroup $H_i < G$, and since $a_0 = 1$, (5.10) becomes

$$1_G = 1_N^G - \sum_{i>0} a_i \lambda_i^G.$$

Since H_i is proper, we may assume that λ_i is an integral linear combination of induced linear characters from elementary subgroups of H_i , and thus $1_G \in \mathcal{J}$ by (4.3)(c). $\square$

(5.8) is a basic result with many important consequences. Here is one interesting one, originally proved by Brauer using block theory.

5.11 (Brauer). *Suppose that $\chi \in \text{Irr}(G)$ and p is a prime. Then the following statements are equivalent:*

- (a) $\chi(g) = 0$ for every $g \in G$ whose order is divisible by p .
- (b) $\chi(g) = 0$ for every $g \in G$ whose order is a power of p .
- (c) $\frac{|G|}{\chi(1)} \not\equiv 0 \pmod{p}$.

PROOF. It is obvious that (a) implies (b). To show that (b) implies (c) let P be a Sylow p -subgroup of G . Then

$$(\chi_P, 1_P) = \frac{1}{|P|} \sum_{x \in P} \chi(x) = \frac{\chi(1)}{|P|} \in \mathbf{Z}.$$

Thus $\chi(1)$ is divisible by the full power of p dividing $|G|$ and (c) follows.

The nontrivial implication is (c) $\Rightarrow$ (a). We first argue that

(5.12) Suppose that $H = P \times Q \subseteq G$ where P is a p -group and $p \nmid |Q|$. Then (λ, χ_Q) is divisible by $|P|$ for all $\lambda \in \text{Irr}(Q)$.

Namely, let $n = \frac{|G|}{\chi(1)}$ and let $x \in Q$. Then by (3.2) the quantity

$$\frac{|G|\chi(x)}{|\mathbf{C}_G(x)|\chi(1)}$$

is an algebraic integer. Since $P \subseteq \mathbf{C}_G(x)$, the quantity

$$\frac{|G|\chi(x)}{|P|\chi(1)} = \frac{n\chi(x)}{|P|}$$

is also an algebraic integer. By hypothesis, there are integers a and b such that $an + b|P| = 1$. Then

$$\frac{\chi(x)}{|P|} = \frac{an\chi(x)}{|P|} + b\chi(x),$$

so $\frac{\chi(x)}{|P|}$ is an algebraic integer as well.

Now choose integers d, e such that $d|P| + e|Q| = 1$. Then

$$\frac{(\chi_Q, \lambda)}{|P|} = d(\chi_Q, \lambda) + \frac{e|Q|}{|P|}(\chi_Q, \lambda) = d(\chi_Q, \lambda) + e \sum_{x \in Q} \frac{\chi(x)}{|P|} \lambda(x^{-1})$$

and since the right-hand side is an algebraic integer, (5.12) follows.

Next, we define a class function $\hat{\chi}$ on G as follows:

$$\hat{\chi}(g) = \begin{cases} \chi(g) & \text{if the order of } g \text{ is not divisible by } p, \\ 0 & \text{otherwise.} \end{cases}$$

We want to prove that $\chi = \hat{\chi}$. The main step is to show, using (5.8), that $\hat{\chi} \in \text{Ch}(G)$. To do that, we only need to show that $\hat{\chi}_H \in \text{Ch}(H)$ for every elementary subgroup H of G . Since cyclic groups are direct products of cyclic groups of prime power order, H is of the form $P \times Q$ where P is a p -group and $|Q| \not\equiv 0 \pmod{p}$. (Q may not be cyclic, but we do not care.) By (2.10) every irreducible character of H is of the form $\chi_{\psi\lambda}$ where $\psi \in \text{Irr}(P)$ and $\lambda \in \text{Irr}(Q)$. Since $\hat{\chi}$ vanishes on elements of order divisible by p , we have

$$(\chi_{\psi\lambda}, \hat{\chi}_H) = \frac{1}{|H|} \sum_{\substack{x \in Q \\ y \in P}} \overline{\psi(y)\lambda(x)} \hat{\chi}(yx) = \frac{1}{|P||Q|} \sum_{x \in Q} \psi(1) \overline{\lambda(x)} \hat{\chi}(x) = \psi(1) \frac{(\lambda, \chi_Q)}{|P|}$$

and hence $\hat{\chi}_H \in \text{Ch}(H)$ by (5.12). By (5.8), $\hat{\chi} \in \text{Ch}(G)$.

Finally, let R be the set of elements of G of order not divisible by p . Then since $1 \in R$, we have

$$0 < (\hat{\chi}, \hat{\chi}) = \frac{1}{|G|} \sum_{g \in R} |\chi(g)|^2 \leq \frac{1}{|G|} \sum_{g \in G} |\chi(g)|^2 = (\chi, \chi) = 1.$$

But since $\hat{\chi}$ is a generalized character, the inequality must be an equality and therefore χ vanishes off R . $\square$

Part II

CHAPTER 6

Permutations and Partitions

In this section, we collect some combinatorial results and introduce some notation which we shall need later. We denote the set consisting of the first n positive integers by Ω^n , and we let S^n be the group of all permutations on this set. We will often omit the superscript n if no confusion seems likely. In a slight departure from usual terminology, we will mean by a *partition* of Ω an ordered collection of pairwise disjoint nonempty subsets $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_r\}$ such that

$$|\mathcal{P}_i| \geq |\mathcal{P}_{i+1}| \quad \text{for all } i, \text{ and } \bigcup_i \mathcal{P}_i = \Omega.$$

(Some authors use the term “tabloid” instead.) Thus, the two partitions

$$(\{1, 2\}, \{3, 4\}) \quad \text{and} \quad (\{3, 4\}, \{1, 2\})$$

of Ω^4 are different. The sets $\mathcal{P}_i$ are the *parts* of $\mathcal{P}$. A partition of Ω is just a surjective function $\Omega \rightarrow \Omega^r$ for some r whose fibers are monotonically decreasing in size.

By a *partition of n* we mean a sequence of positive integers

$$\pi = (\pi_1, \pi_2, \dots, \pi_r)$$

such that

$$\pi_i \geq \pi_{i+1} \quad \text{for all } i \text{ and } \sum_{i=1}^r \pi_i = n.$$

The integers π_i are the *parts* of π . We often indicate repeating terms exponentially, so $(3^2, 2, 1^3)$ means $(3, 3, 2, 1, 1, 1)$.

Given a partition $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_r)$ of Ω we let the *type* of $\mathcal{P}$ be the partition $\overline{\mathcal{P}} = (\pi_1, \pi_2, \dots, \pi_r)$ where $\pi_i = |\mathcal{P}_i|$ for all i . We will sometimes abuse notation by considering partitions of Ω (resp. of n) as infinite sequences whose parts are eventually empty (resp. zero). Furthermore, it is sometimes convenient to drop the restriction that the parts of a partition are monotonically decreasing. When we wish to relax this condition, we shall call the partition *improper*.

The symmetric group S^n acts on the set of all partitions of Ω^n in an obvious way. This action evidently preserves types, and if two partitions have the same type, it is clear that we can relabel the elements of one to obtain the other. The stabilizer $S_{\mathcal{P}}$ of a partition $\mathcal{P}$ is called a *Young subgroup of type $\mathcal{P}$* . If $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_r)$ and we let S_i be the subgroup of S fixing $\Omega \setminus \mathcal{P}_i$, then $S_{\mathcal{P}} \cong S_1 \times S_2 \times \dots \times S_r$. Suppose that $\mathcal{Q} = (\mathcal{Q}_1, \dots, \mathcal{Q}_s)$ is another partition.

If we let $\mathcal{R}$ be a partition whose parts are the sets $\mathcal{P}_i \cap \mathcal{Q}_j$ in an appropriate order, then $S_{\mathcal{P}} \cap S_{\mathcal{Q}} = S_{\mathcal{R}}$. To summarize:

(6.1) *Two partitions of Ω are S -conjugate if and only if they have the same type. If $S_{\mathcal{P}}$ and $S_{\mathcal{Q}}$ are Young subgroups, then $S_{\mathcal{P}} \cap S_{\mathcal{Q}}$ is the Young subgroup of a partition whose parts are the nonempty pairwise intersections of a part of $\mathcal{P}$ with a part of $\mathcal{Q}$. $\square$*

To each partition $\pi = (\pi_1, \pi_2, \dots, \pi_r)$ of n we associate a *Young diagram* $Y(\pi) = \{(i, j) : 1 \leq i \leq r, 1 \leq j \leq \pi_i\}$. We often think of Young diagrams as arrays of boxes, for example,

$$Y(3^2, 2, 1) = \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \\ \hline \square & & \\ \hline \end{array}$$

Given a Young diagram $Y(\pi)$, it is not difficult to see that the transpose diagram $Y(\pi)' = \{(j, i) \mid (i, j) \in Y(\pi)\}$ is the Young diagram of a uniquely determined *opposite partition*, π' . For example, $(3^2, 2, 1)' = (4, 3, 2)$. Some authors call π' the *conjugate partition* to π .

Let π be a partition of n . A *Young tableau* of type π is a bijection $T : Y(\pi) \rightarrow \Omega^n$. This can be thought of as an assignment of numbers to boxes. The following is a tableau of type $(3^2, 2, 1)$:

5	1	8
3	9	2
6	7	
4		

Each tableau also has an opposite tableau. Moreover, any tableau T defines two partitions of Ω , the *row partition* $\mathcal{R}(T)$ and the *column partition* $\mathcal{C}(T)$. Any two partitions which are obtained from a single Young diagram in this way will be called *opposite*. More generally, we say that $\mathcal{P}$ is *disjoint from* $\mathcal{Q}$ if $|\mathcal{P}_i \cap \mathcal{Q}_j| \leq 1$ for every part $\mathcal{P}_i$ of $\mathcal{P}$ and $\mathcal{Q}_j$ of $\mathcal{Q}$. It is clear that S acts freely on the set of tableaux of a given type by permuting the entries. We define the *row group* $R(T)$ (resp. *column group* $C(T)$) to be the stabilizer of the row partition (resp. column partition) of T .

(6.2) *Let $\mathcal{P}$ and $\mathcal{Q}$ be partitions of Ω . Then $\mathcal{P}$ and $\mathcal{Q}$ are opposite iff they are disjoint and have opposite types. Moreover, $S_{\mathcal{P}}$ is transitive on the set of partitions opposite to $\mathcal{P}$.*

PROOF. It is clear from the definition that opposite partitions have opposite types and are disjoint. Conversely, if $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_r)$ and $\mathcal{Q} = (\mathcal{Q}_1, \dots, \mathcal{Q}_s)$ are disjoint and have opposite types, we must construct a tableau T of type $\overline{\mathcal{P}}$ with $\mathcal{R}(T) = \mathcal{P}$ and $\mathcal{C}(T) = \mathcal{Q}$. Since $\overline{\mathcal{P}} = \overline{\mathcal{Q}}'$, it follows in particular that $|\mathcal{Q}_1| = r$, and since $|\mathcal{Q}_1 \cap \mathcal{P}_i| \leq 1$ for $1 \leq i \leq r$, we conclude

that $|\mathcal{Q}_1 \cap \mathcal{P}_i| = 1$ for all i . Thus, we can let $T(i, 1)$ be the unique element of $\mathcal{Q}_1 \cap \mathcal{P}_i$ for each i . We can now fill in the remaining columns of T inductively by setting $\tilde{\Omega} = \Omega \setminus \mathcal{Q}_1$, $\tilde{\mathcal{P}}_i = \mathcal{P}_i \setminus \mathcal{Q}_1$ ($1 \leq i \leq r$), and $\tilde{\mathcal{Q}}_i = \mathcal{Q}_{i+1}$ ($1 \leq i < s$).

Finally, given two tableaux T and $\tilde{T}$ with the same row partition, it is obvious that one can be obtained from the other by permuting the elements in each row. Thus, there is an element $\sigma \in R(T)$ such that $\mathcal{E}(T)^\sigma = \mathcal{E}(\tilde{T})$. $\square$

Given any partition $\mathcal{P}$ of Ω , there is a standard way to choose a tableau T with $\mathcal{R}(T) = \mathcal{P}$, namely arrange each part of $\mathcal{P}$ in monotonically decreasing order. The resulting tableau is row-monotonic. Tableaux which are both row- and column-monotonic are called *standard*.

We next introduce an important partial order on (improper) partitions of n , defined as follows:

$$(\lambda_1, \dots, \lambda_r) \ll (\mu_1, \dots, \mu_s) \text{ iff } \sum_{i=1}^k \lambda_i \leq \sum_{i=1}^k \mu_i \text{ for } 1 \leq k \leq n,$$

where we are taking $\lambda_i = \mu_j = 0$ for $i > r$ and $j > s$. Denote by $\leq$ the following total lexicographic order:

$$\lambda \leq \mu \text{ iff } \lambda = \mu \text{ or for some } k \text{ we have } \lambda_i = \mu_i \text{ for } i < k \text{ and } \lambda_k < \mu_k.$$

It is clear that if $\lambda \leq \mu$ then $\lambda \ll \mu$. The next result gives an elegant and important characterization of disjointness in terms of this partial order. We only need the easy implication, but prove both for the sake of completeness.

6.3 (Gale-Ryser). *Let λ and μ be partitions of n . There exist disjoint partitions of Ω^n of types λ and μ iff $\lambda \ll \mu'$.*

PROOF. Here is a good way to think about this result: The rows of $Y(\lambda)$ are families who are going on a bus trip, each box denoting a family member. The rows of $Y(\mu)$ are the buses, each box denoting a seat. We are looking for a “harmonious” seating arrangement, that is, one in which no two family members are seated on the same bus.

There is an obvious necessary condition provided by the pigeonhole principle, namely that after the first k families are seated there be no more than k persons per bus. Let $C_k = \sum_{j=1}^k \mu'_j$. Then C_k is the total number of boxes in the first k columns of $Y(\mu)$, which is the total number of seats available subject to the constraint that there be no more than k persons per bus. We will call C_k the “ k -capacity” of the buses. It must be at least as large as the total size of the largest k families. It follows that if disjoint partitions exist, then $\lambda \ll \mu'$.

Conversely, we assume that the total size of the largest k families does not exceed the k -capacity of the buses for any k . We put as many people as can be seated harmoniously (e.g., at most one from each family) on the largest bus, send it on its way, and proceed by induction. The problem is to verify that the remaining people and buses satisfy the constraint that the new total size of the largest k families does not exceed the k -capacity of the remaining buses.

Let s be the size of the bus just dispatched, then the k -capacity of the remaining fleet has been reduced by k for all $k \leq s$, and by s for all $k \geq s$. On the other hand, since at most one person has been removed from each family, the total size of the largest k families has been reduced by at most k for all $k \leq s$, and by at most s for any k , since at most s people left on

the first bus. It follows by induction that the remaining people can be seated harmoniously. $\square$

One interesting corollary of (6.3) is that the relation $\lambda \ll \mu'$ is symmetric. More importantly, however, recall from (6.1) that two Young subgroups have trivial intersection iff they are the stabilizers of disjoint partitions. Thus, we have

(6.4) COROLLARY. *Let λ and μ be partitions of n . There exist Young subgroups of types λ and μ with trivial intersection iff $\lambda \ll \mu'$. $\square$*

For $\sigma \in S^n$ let $\langle \sigma \rangle$ be the cyclic subgroup generated by σ and let

$$\Omega = \Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_r$$

where the Ω_i are the (disjoint) orbits of $\langle \sigma \rangle$ on Ω . Let $|\Omega_i| = k_i$ ($1 \leq i \leq r$) with notation chosen so that $k_i \geq k_{i+1}$. We call the partition $\bar{\sigma} = (k_1, \dots, k_r)$ the *type* of σ . We say that σ is a *k-cycle* if $k_1 = k$ and $k_2 = 1$. The usual notation for a *k-cycle* σ is $(m_0 m_1 \cdots m_{k-1})$ where $m_i^\sigma = m_{i+1}$ ($0 \leq i < k$). This notation is unique up to a cyclic permutation of the m_i . Moreover, it has the further advantage that $\tau^{-1} \sigma \tau = (m_0^\tau m_1^\tau \cdots m_{k-1}^\tau)$, whence it is obvious that any two *k-cycles* are S^n -conjugate.

Returning to the general case, let σ_i be the k_i -cycle which agrees with σ on Ω_i and is the identity elsewhere. Then the σ_i are *disjoint* (meaning that their nontrivial orbits on Ω are disjoint) and their product is σ . It is easy to see that the σ_i are uniquely determined by σ , thus there is a unique way of writing σ as a product of disjoint cycles, up to the order of the factors (which is irrelevant since the σ_i obviously commute). Furthermore, it is clear that σ is S -conjugate to τ iff $\bar{\sigma} = \bar{\tau}$. To summarize:

(6.5) *Every element of S^n is uniquely the product of disjoint cycles. The lengths of these cycles form a partition of n and, in this way, the conjugacy classes of S^n are indexed by the partitions of n . $\square$*

Since the number of conjugacy classes equals the number of irreducible characters, we might hope that there is also a natural way to index the irreducible characters of S^n by the partitions of n . This indeed turns out to be the case, as we shall see in the next section.

For computational purposes, it is important to know the order of each conjugacy class in S^n , or what is the same thing, the order of the centralizer of a representative element.

(6.6) *For any partition $\pi = n^{j_n} \cdots 2^{j_2} 1^{j_1}$ in exponential form, define*

$$n_\pi = \prod_{i=1}^n j_i! i^{j_i}.$$

Then $|\mathbf{C}_{S^n}(\sigma)| = n_{\bar{\sigma}}$ for any $\sigma \in S^n$.

PROOF. For $i = 1, 2, \dots, n$, let m_i be the number of orbits of σ of size i and let $\mathcal{Q}_i$ be the union of these orbits; then $|\mathcal{Q}_i| = im_i$. Let H_i be the subgroup of S^n which permutes these orbits and is the identity off $\mathcal{Q}_i$. H_i has a normal subgroup N_i which stabilizes each of the m_i orbits. N_i is a direct product of m_i copies of S^i , and $H_i/N_i \cong S^{m_i}$. Let $C_i = \mathbf{C}_{H_i}(\sigma)$. By (6.5)

we can write $\sigma = \sigma_1 \sigma_2 \cdots \sigma_n$, where for each i , $\sigma_i \in N_i$ is a product of m_i disjoint i -cycles (with the convention that the empty product is the identity). Our formula is a consequence of the following three facts, each of which is fairly obvious:

- (i) $\mathbf{C}_{S^n}(\sigma) = C_1 \times C_2 \times \cdots \times C_n$,
- (ii) $C_i N_i = H_i$, and
- (iii) $C_i \cap N_i = \mathbf{C}_{N_i}(\sigma_i) = \langle \sigma_i \rangle$.

Assertion (i) is just a consequence of the fact that $\mathbf{C}_{S^n}(\sigma)$ must permute orbits of σ of the same size. Assertion (ii) follows, for example, by constructing a product of transpositions which interchanges two orbits of σ and is the identity elsewhere. Assertion (iii) quickly reduces to the statement that the only elements of S^i commuting with an i -cycle are its powers, a fact whose proof is an easy exercise. $\square$

CHAPTER 7

The Irreducible Characters of S^n

In this section, we will define, for each partition π of n , an irreducible character χ_π of S^n , and describe an effective algorithm for computing the character table of S^n .

First, recall that S^n acts on the ring of polynomials in n commuting variables $x_1, \dots, x_n$ by permuting indices. Let $\Delta = \prod_{i < j} (x_i - x_j)$ and let $\sigma \in S^n$. Then $\sigma(\Delta) = \text{sgn}(\sigma)\Delta$ where $\text{sgn} : S^n \rightarrow \{\pm 1\}$ is a linear character of S^n , called the *signature*. We will often use the notation $(-1)^\sigma = \text{sgn}(\sigma)$. A permutation σ is *odd* if $(-1)^\sigma = -1$ and *even* if $(-1)^\sigma = +1$. Note that $(-1)^{(1,2)} = -1$.

Now consider the action of the symmetric group on the set of partitions of Ω . As we remarked in the previous section, S is transitive on partitions of a given type. Let S_π be a Young subgroup of type π , and let $\psi_\pi = 1_{S_\pi}^S$, the permutation character of S afforded by the action on partitions of type π for any partition π of n . In addition, let $\phi_\pi = (-1)_{S_\pi}^S$, the signature character of S_π induced to S .

(7.1) $(\phi_\mu, \psi_\lambda) \neq 0$ iff $\mu' \gg \lambda$. Moreover, if $\mu' = \lambda$ then $(\phi_\mu, \psi_\lambda) = 1$.

PROOF. Let K and H be Young subgroups of types λ and μ respectively, chosen with $K \cap H = 1$ if possible. By (4.4) we have

$$(\phi_\mu, \psi_\lambda) = \sum_{i=1}^t (1_{H_i}, (-1)_{H_i}^{(i)})$$

where $H_i = H \cap x_i^{-1} K x_i$ and the x_i are (K, H) double coset representatives. Since the signature is constant on S -conjugacy classes, $(-1)^{(i)} = -1$. But H_i is an intersection of Young subgroups of types μ and λ . If $H_i \neq 1$, then H_i contains an odd permutation by (6.1), whence 1_{H_i} and -1_{H_i} are distinct irreducible characters of H_i whose inner product is therefore zero. Hence, $(\phi_\mu, \psi_\lambda) \neq 0$ iff $H_i = 1$ for some i , and the first assertion follows from (6.4).

Now assume $\mu' = \lambda$. Then what we must prove is that there is exactly one value of i for which $H_i = 1$. Choose disjoint partitions $\mathcal{P}$ and $\mathcal{Q}$ of types μ and λ respectively and take $H = S_{\mathcal{P}}$ and $K = S_{\mathcal{Q}}$. Choose notation so that $x_1 = 1$, then $H_1 = K \cap H = 1$. Suppose that for some $\sigma \in S$, $K^\sigma \cap H = 1$. Then $\mathcal{Q}^\sigma$ is disjoint from $\mathcal{P}$ by (6.1) and therefore $\mathcal{Q}^{\sigma h} = \mathcal{P}$ for some $h \in H$ by (6.2). Thus $\sigma \in KH$ as required. $\square$

The previous result is critical. It says that for any partition π of n , ψ_π and $\phi_{\pi'}$ have a unique common irreducible constituent, which moreover has

multiplicity one in each of them. We now define χ_π to be this unique common irreducible constituent of ψ_π and $\phi_{\pi'}$.

For each partition λ of n , let $\sigma_\lambda \in S$ be a conjugacy class representative, and let $\mathbf{X} = \chi_\pi(\sigma_\lambda)$ and $\mathbf{Y} = \psi_\pi(\sigma_\lambda)$ be square matrices indexed by the partitions of n . Order the rows of $\mathbf{X}$ and $\mathbf{Y}$ in descending lexicographic order (with (n) first and (1^n) last) and order the columns in ascending lexicographic order.

Suppose $(\psi_\mu, \chi_\lambda) \neq 0$. Then since χ_λ is by definition a constituent of $\phi_{\lambda'}$, we certainly have $(\psi_\mu, \phi_{\lambda'}) \neq 0$ and hence $\lambda \gg \mu$ by (7.1). Suppose that $\chi_\lambda = \chi_{\hat{\lambda}}$ for some partition $\hat{\lambda}$. Then $(\psi_\lambda, \chi_{\hat{\lambda}}) = 1 = (\psi_{\hat{\lambda}}, \chi_{\hat{\lambda}})$ whence $\lambda \gg \hat{\lambda} \gg \lambda$ and thus $\lambda = \hat{\lambda}$. The χ_λ are therefore distinct irreducible characters, and since $\mathbf{X}$ is square, it must be the character table of S . Since the lexicographic order $\geq$ is a refinement of the partial order $\gg$, we see that $\mathbf{Y} = \mathbf{LX}$ for some lower triangular integral matrix $\mathbf{L}$ with ones on the diagonal. In particular, $\mathbf{L}$ is invertible over $\mathbf{Z}$. Thus, we have proved

(7.2) *If $(\psi_\mu, \chi_\lambda) \neq 0$ then $\lambda \gg \mu$. In particular, the χ_λ are distinct, $\mathbf{X}$ is the character table of S , and $\mathbf{Y} = \mathbf{LX}$ where $\mathbf{L}$ is a lower-triangular matrix with ones on the diagonal. Moreover, the ψ_μ are a $\mathbf{Z}$ -basis for the space of generalized characters of S . $\square$*

We can now describe a very simple recursive algorithm for the computation of $\mathbf{X}$. Initially, we have $\chi_{(n)} = \psi_{(n)} = 1_S$. Now assume that we have computed ψ_μ and that we have already computed χ_λ for all $\lambda > \mu$. Then

$$\chi_\mu = \psi_\mu - \sum_{\lambda > \mu} (\psi_\mu, \chi_\lambda) \chi_\lambda.$$

Thus, the μ th row of $\mathbf{X}$ is computed by first taking inner products of the μ th row of $\mathbf{Y}$ with all previous rows of $\mathbf{X}$ and then subtracting the appropriate multiples of the previous rows of $\mathbf{X}$ from the μ th row of $\mathbf{Y}$. In order to do this, we need to know how to compute $\mathbf{Y}$, but this is relatively straightforward. Namely, given two partitions $\pi = (\pi_1, \dots, \pi_r)$ and $\lambda = (\lambda_1, \dots, \lambda_s)$ of n , we define a λ -refinement of π to be a surjective function $f: \Omega^\pi \rightarrow \Omega^\lambda$ such that $\pi_j = \sum_{f(i)=j} \lambda_i$ ($1 \leq j \leq r$).

(7.3) *Let π and λ be partitions of n . Then $\psi_\pi(\sigma_\lambda)$ is the number of λ -refinements of π .*

PROOF. We count the number of partitions of Ω of type π which are fixed by an element of type λ . Let σ be of type $\lambda = (\lambda_1, \dots, \lambda_s)$ and let $\sigma = \sigma_1 \sigma_2 \cdots \sigma_s$ be the decomposition of σ as a product of disjoint cycles, where σ_i is a λ_i -cycle. Let $\mathcal{S} = (\mathcal{S}_1, \dots, \mathcal{S}_s)$ be the corresponding partition of Ω , so that σ_i permutes $\mathcal{S}_i$ cyclically and fixes the remaining points of Ω . In order that σ fix a partition $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_r)$, it is necessary and sufficient that each $\mathcal{S}_i$ be contained in some $\mathcal{P}_j$. If this happens for some partition $\mathcal{P}$ of type π , set $f(i) = j$ to obtain a function f which is easily seen to be a λ -refinement of π . Conversely, given such a function f , let $\mathcal{P}_j = \bigcup_{f(i)=j} \mathcal{S}_i$ for each j to obtain a partition $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_r)$ of type π fixed by σ . $\square$

Observe that $\phi_\pi = (-1) \cdot \psi_\pi$ by (4.3)(a). Since χ_π is the unique irreducible constituent of ψ_π and $\phi_{\pi'}$, it follows that $(-1) \cdot \chi_\pi$ must be the unique irreducible constituent of $(-1) \cdot \psi_\pi$ and $(-1) \cdot \phi_{\pi'}$. Since $(-1) \cdot \psi_\pi = \phi_\pi$ and $(-1) \cdot \phi_{\pi'} = \psi_{\pi'}$, we have proved

$$(7.4) \quad \chi_{\pi'} = (-1) \cdot \chi_\pi. \quad \square$$

We can use the above results to compute the character table of S^6 . By (7.4), we only need to compute one of $\{\chi_\pi, \chi_{\pi'}\}$ which saves approximately half the work. We first compute $\mathbf{Y}$:

class	1^6	$2^1 1^4$	$2^2 1^2$	2^3	$3^1 1^3$	$3^1 2^1 1^1$	3^2	$4^1 1^2$	$4^1 2^1$	$5^1 1^1$	6^1
#elts	1	15	45	15	40	120	40	90	90	144	120
6^1	1	1	1	1	1	1	1	1	1	1	1
$5^1 1^1$	6	4	2	0	3	1	0	2	0	1	0
$4^1 2^1$	15	7	3	3	3	1	0	1	1	0	0
$4^1 1^2$	30	12	2	0	6	0	0	2	0	0	0
3^2	20	8	4	0	2	2	2	0	0	0	0
$3^1 2^1 1^1$	60	16	4	0	3	1	0	0	0	0	0

As an example of how (7.3) is used to compute $\mathbf{Y}$, consider the calculation $\psi_{4^1 1^2}(\sigma_{2^1 1^4}) = 12$. We are counting the number of ways that the parts of $4^1 1^2$ can be written as sums of parts of $2^1 1^4$. Clearly, 2 must be a summand of 4 along with two other 1's, so there are six different ways to write 4 as a sum of parts of $2^1 1^4$. Having chosen one such way, each of the remaining 1's is just the sum of one of the remaining 1's of $2^1 1^4$, so there are two ways to do this. Hence the number of $2^1 1^4$ -refinements of $4^1 1^2$ is 12.

Now given $\mathbf{Y}$, the computation of $\mathbf{X}$ is completely mechanical. The first row is all 1's. To get the second row, we compute the inner product of the first row of $\mathbf{X}$ with the second row of $\mathbf{Y}$ to get the multiplicity of χ_{6^1} in $\psi_{5^1 1^1}$ (which is of course 1) and subtract that multiple of the first row of $\mathbf{X}$ from the second row of $\mathbf{Y}$ to get the second row of $\mathbf{X}$. For the third row of $\mathbf{X}$, we compute the inner products of the first two rows of $\mathbf{X}$ with the third row of $\mathbf{Y}$ (which are both 1), and subtract from the third row of $\mathbf{Y}$ those multiples of the first two rows of $\mathbf{X}$. Notice that at each stage, we can check our work by computing the inner product of each row of $\mathbf{X}$ with itself (it must be 1). The results are as follows:

class	1^6	$2^1 1^4$	$2^2 1^2$	2^3	$3^1 1^3$	$3^1 2^1 1^1$	3^2	$4^1 1^2$	$4^1 2^1$	$5^1 1^1$	6^1
#elts	1	15	45	15	40	120	40	90	90	144	120
6^1	1	1	1	1	1	1	1	1	1	1	1
$5^1 1^1$	5	3	1	-1	2	0	-1	1	-1	0	-1
$4^1 2^1$	9	3	1	3	0	0	0	-1	1	-1	0
$4^1 1^2$	10	2	-2	-2	1	-1	1	0	0	0	0
3^2	5	1	1	-3	-1	1	2	-1	-1	-1	0
$3^1 2^1 1^1$	16	0	0	0	-2	0	-2	0	0	0	1

CHAPTER 8

The Specht Modules

We now turn to the problem of constructing a module X_π affording the irreducible character χ_π . Our treatment here follows James [4]. We begin by letting M_π be the permutation module affording ψ_π . M_π has a natural basis $\{f_{\mathcal{P}} \mid \overline{\mathcal{P}} = \pi\}$ permuted by S . Since χ_π has multiplicity 1 in ψ_π , there is a unique submodule $X_\pi \subseteq M_\pi$ affording χ_π which is called the *Specht module*. We want to construct an explicit basis for X_π in terms of the $f_{\mathcal{P}}$. Given any partition $\mathcal{Q}$, we define

$$\tau_{\mathcal{Q}} = \sum_{g \in S_{\mathcal{Q}}} (-1)^g g \in \mathbb{C}G.$$

(8.1) *Let $\mathcal{Q}$ be any partition of Ω of type π' . Then $M_\pi \tau_{\mathcal{Q}} \subseteq X_\pi$.*

PROOF. Let H be a Young subgroup of type π , so that $\psi_\pi = 1_H^G$. Let

$$e_H = \frac{1}{|H|} \sum_{h \in H} h;$$

then e_H is the primitive central idempotent of $\mathbb{C}H$ corresponding to 1_H . By (4.9), the right ideal $I_\pi = e_H \mathbb{C}G$ affords ψ_π and is thus isomorphic to M_π . Since $\frac{1}{|S_{\mathcal{Q}}|} \tau_{\mathcal{Q}}$ is the primitive central idempotent of $\mathbb{C}S_{\mathcal{Q}}$ corresponding to $(-1)_H$, the right ideal $J_{\pi'} = \tau_{\mathcal{Q}} \mathbb{C}G$ affords $\phi_{\pi'}$.

Let B_χ be the minimal 2-sided ideal of $\mathbb{C}G$ corresponding to an irreducible character χ of G . Then $I_\pi B_\chi \subseteq I_\pi \cap B_\chi$. Since every irreducible submodule of B_χ affords χ by (1.3), $I_\pi B_\chi = 0$ unless $(\psi_\pi, \chi) \neq 0$. Similarly, $B_\chi J_{\pi'} = 0$ unless $(\chi, \phi_{\pi'}) \neq 0$. Since $\mathbb{C}G$ is the sum of its minimal 2-sided ideals, we have first that

$$I_\pi \subseteq \sum_{(\chi, \psi_\pi) \neq 0} B_\chi,$$

and then that $I_\pi J_{\pi'} \subseteq B_{\chi_\pi}$ because ψ_π and $\phi_{\pi'}$ have a unique irreducible constituent, namely χ_π , in common. We conclude that $M_\pi \tau_{\mathcal{Q}}$ is contained in a submodule of M_π all of whose irreducible constituents afford χ_π . $\square$

Next, given any tableau T of type π with $\mathcal{R} = \mathcal{R}(T)$ and $\mathcal{C} = \mathcal{C}(T)$, we define the *Specht vector* $v(T) = f_{\mathcal{R}} \tau_{\mathcal{C}}$. By (8.1), $v(T) \in X_\pi$. If T is a standard tableau, we call $v(T)$ a *standard Specht vector*. For bookkeeping purposes in the proof of the next result, it is convenient to introduce the following total ordering. Given two partitions $\mathcal{P}$ and $\mathcal{Q}$ of Ω of the same type with associated surjective functions $p, q : \Omega \rightarrow \Omega'$ we define $\mathcal{P} \leq \mathcal{Q}$ if $\mathcal{P} = \mathcal{Q}$ or

there is an i such that $p(j) = q(j)$ for $j > i$ and $p(i) > q(i)$. This is evidently a total order on the set of partitions of a given type. Now given two tableaux T, T' of the same type, define $T > T'$ if $\mathcal{C}(T) > \mathcal{C}(T')$, or if $\mathcal{C}(T) = \mathcal{C}(T')$ and $\mathcal{R}(T) > \mathcal{R}(T')$.

(8.2) *The set of all Specht vectors of type π is permuted by S , and the subset of standard Specht vectors spans X_π .*

PROOF. Let $C = C(T)$ be the column group of T and let $\mathcal{C} = \mathcal{C}(T)$ and $\mathcal{R} = \mathcal{R}(T)$ be the column and row partitions of T respectively. For any $\sigma \in S$, we have

$$C^\sigma = \sigma^{-1} C \sigma = C(T^\sigma)$$

and

$$\sigma^{-1} \tau_{\mathcal{C}} \sigma = \sum_{g \in C} (-1)^g \sigma^{-1} g \sigma = \sum_{g \in C^\sigma} (-1)^g g = \tau_{\mathcal{C}^\sigma}.$$

Hence,

$$v(T) \cdot \sigma = f_{\mathcal{R}} \cdot \sigma \sigma^{-1} \cdot \tau_{\mathcal{C}} \cdot \sigma = f_{\mathcal{R}} \cdot \sigma \cdot \tau_{\mathcal{C}^\sigma} = f_{\mathcal{R}^\sigma} \cdot \tau_{\mathcal{C}^\sigma} = v(T^\sigma).$$

Since the $v(T)$ are permuted by S , they must span an S -submodule of X_π , but since X_π is irreducible, they span X_π .

In order to show that the standard $v(T)$ span X_π , it suffices to show that if T is not standard then there exist integers $a_{T'}$ such that

$$(8.3) \quad v(T) = \sum_{T' > T} a_{T'} v(T').$$

In fact, we will show that the nonzero $a_{T'}$ can be chosen to be ± 1 . We first observe that if $\sigma \in C(T)$ then $\tau_{\mathcal{C}^\sigma} = (-1)^\sigma \tau_{\mathcal{C}}$ and thus

$$(8.4) \quad v(T^\sigma) = v(T) \sigma = (-1)^\sigma v(T).$$

It is clear that if T is column monotonic then $T \geq T^\sigma$ for any $\sigma \in C(T)$, hence using (8.4) if necessary, we may assume that T is column monotonic. Since T is not standard, we have $T(i, j) < T(i, k)$ for some $j < k$ and some i . Let n_j and n_k be the lengths of columns j and k of T , and consider the subset

$$\begin{aligned} \Omega_0 = \{ & T(1, k) > T(2, k) > \cdots > T(i, k) > T(i, j) \\ & > T(i+1, j) > \cdots > T(n_j, j) \} \end{aligned}$$

of Ω . Let H be the subgroup of S which is the identity off Ω_0 , and let $\tau_H = \sum_{h \in H} (-1)^h h$. We claim that $v(T) \tau_H = 0$. Since $v(T)$ is an alternating sum of standard basis vectors $f_{\mathcal{R}^\sigma}$ as σ ranges over $C(T)$, it suffices to show that $f_{\mathcal{R}^\sigma} \tau_H = 0$ for all $\sigma \in C(T)$.

To see this, choose $\sigma \in C(T)$, and note first that $|\Omega_0| = n_j + 1$, and since $n_j \geq n_k$, there will always be at least one row m of T^σ such that $\alpha = T^\sigma(m, j)$ and $\beta = T^\sigma(m, k)$ are both elements of Ω_0 . Let $h_0 \in H$ be the transposition (α, β) and let $\{h_1, \dots, h_t\}$ be right coset representatives for $\langle h_0 \rangle$ in H . Then $(-1)^{h_0 h_i} = -(-1)^{h_i}$ and $f_{\mathcal{R}^\sigma} h_0 = f_{\mathcal{R}^\sigma}$, whence

$$f_{\mathcal{R}^\sigma} \tau_H = \sum_{i=1}^t (-1)^{h_i} f_{\mathcal{R}^\sigma} h_i + \sum_{i=1}^t (-1)^{h_0 h_i} f_{\mathcal{R}^\sigma} h_0 h_i = 0.$$

We have therefore derived the relation

$$(8.5) \quad \sum_{h \in H} (-1)^h v(T^h) = 0.$$

However, (8.5) has many repeated terms. To collect them, let $H_0 = H \cap C(T)$ and let X be a set of representatives for the nonidentity right cosets of H_0 in H . Since $v(T^{hx}) = (-1)^h v(T^x)$ for $h \in H_0$, (8.5) implies

$$(8.6) \quad v(T) + \sum_{x \in X} (-1)^x v(T^x) = 0.$$

Since $X \cap H_0 = \emptyset$, each element of X moves at least one element of column k to column j . But every element of Ω_0 in column k is bigger than every element of Ω_0 in column j , so that the largest element of T which is moved by any $x \in X$ is moved to a lower-numbered column. It follows that $T^x > T$ and thus (8.6) is of the form (8.3), and the proof of (8.2) is complete. $\square$

(8.7) *If $T_1, \dots, T_t$ are standard tableaux with $\mathcal{R}(T_1) > \mathcal{R}(T_i)$ for $i > 1$, and $\sum_{i=1}^t a_i v(T_i) = \sum_{\mathcal{P}} b_{\mathcal{P}} f_{\mathcal{P}}$, then $b_{\mathcal{P}} = 0$ for $\mathcal{P} > \mathcal{R}(T_1)$ and $b_{\mathcal{R}(T_1)} = a_1$. In particular, the standard Specht vectors of type π are a basis for X_{π} , and $\chi_{\pi}(1)$ is the number of standard tableaux of type π .*

PROOF. By the definition of $v(T)$, we have

$$(8.8) \quad \sum_{i=1}^t a_i v(T_i) = \sum_{i=1}^t a_i f_{\mathcal{R}_i} \tau_{\mathcal{C}_i} = \sum_{i=1}^t a_i \sum_{g \in C(T_i)} (-1)^g f_{R_i^g}$$

where $\mathcal{R}_i = \mathcal{R}(T_i)$ and $\mathcal{C}_i = \mathcal{C}(T_i)$. But when T is column monotonic we also have $\mathcal{R}(T) > \mathcal{R}(T^\sigma)$ for any nonidentity $\sigma \in C(T)$. This implies that the coefficient of $f_{\mathcal{R}_1}$ in (8.8) is a_1 and that the coefficient of $f_{\mathcal{P}}$ is 0 for $\mathcal{P} > \mathcal{C}_1$.

Now suppose that there is a dependence relation on standard Specht vectors: $\sum_{i=1}^t a_i v(T_i) = 0$. Since any two standard tableaux with the same row partition are equal, notation can be chosen so that $\mathcal{R}_1 > \mathcal{R}_i$ for all $i > 1$. But then $a_1 = 0$ by the above, and hence $a_i = 0$ for all i by an obvious induction argument. $\square$

Notice that relations (8.4) and (8.6) express nonstandard Specht vectors as $\mathbf{Z}$ -linear combinations of standard ones. Since the Specht vectors are permuted by S , it follows that X_{π} is defined over $\mathbf{Z}$ and the standard Specht vectors are a $\mathbf{Z}$ -basis. Notice also that (8.7) provides a constructive algorithm for finding representing matrices. Namely, if $v = \sum_{\mathcal{P}} b_{\mathcal{P}} f_{\mathcal{P}} \in X_{\pi}$ and $\mathcal{Q} = \max\{\mathcal{P} \mid b_{\mathcal{P}} \neq 0\}$, then (8.7) implies that the unique row-monotonic tableau $T_{\mathcal{Q}}$ with $\mathcal{R}(T_{\mathcal{Q}}) = \mathcal{Q}$ is in fact standard, and that $v - b_{\mathcal{Q}} v(T_{\mathcal{Q}})$ is a linear combination of standard Specht vectors which are smaller than $T_{\mathcal{Q}}$.

As an example, consider the case $n = 6$, $\pi = (3^2)$. Since π has just two parts, we can specify a partition of Ω of type π by just giving its first part. The standard tableaux of type (3^2) , listed with row partitions in descending order, are

$$T_1 = \begin{array}{|c|c|c|} \hline 6 & 5 & 4 \\ \hline 3 & 2 & 1 \\ \hline \end{array}, \quad T_2 = \begin{array}{|c|c|c|} \hline 6 & 5 & 3 \\ \hline 4 & 2 & 1 \\ \hline \end{array}, \quad T_3 = \begin{array}{|c|c|c|} \hline 6 & 5 & 2 \\ \hline 4 & 3 & 1 \\ \hline \end{array},$$

$$T_4 = \begin{array}{|c|c|c|} \hline 6 & 4 & 3 \\ \hline 5 & 2 & 1 \\ \hline \end{array}, \quad T_5 = \begin{array}{|c|c|c|} \hline 6 & 4 & 2 \\ \hline 5 & 3 & 1 \\ \hline \end{array}.$$

Each standard Specht vector is an alternating sum of eight of the 20 different basis vectors $f_{\mathcal{P}}$. For instance,

$$v(T_1) = f_{654} - f_{354} - f_{624} - f_{651} + f_{324} + f_{351} + f_{621} - f_{321}.$$

Say we want to calculate the matrix of the transposition $\sigma = (1, 2)$. Then

$$v(T_1)\sigma = f_{654} - f_{354} - f_{614} - f_{652} + f_{314} + f_{352} + f_{621} - f_{321},$$

and since the largest term is f_{654} we should subtract $v(T_1)$:

$$v(T_1)\sigma - v(T_1) = f_{624} - f_{614} + f_{651} - f_{652} + f_{314} - f_{324} + f_{352} - f_{351}.$$

The largest term in this expression is f_{652} , so we should add $v(T_3)$:

$$v(T_3) = f_{652} - f_{452} - f_{632} - f_{651} + f_{432} + f_{451} + f_{631} - f_{431},$$

so remembering that the f subscripts are unordered, we have

$$v(T_1)\sigma - v(T_1) + v(T_3) = f_{624} - f_{614} + f_{352} - f_{351} - f_{452} - f_{632} + f_{451} + f_{631} \\ = v(T_5).$$

In a similar way, one can rewrite $v(T_i)\sigma$ in terms of the $v(T_j)$ for $i = 2, 3, 4, 5$. The resulting matrix is

$$\mathbf{X}_{\pi}(1, 2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 & -1 \end{bmatrix}.$$

We can use the Specht modules to analyze the restriction of an irreducible character χ_{λ} of S^n to S^{n-1} . Let $\lambda = (\lambda_1, \dots, \lambda_r)$ and let $\{i_1, i_2, \dots, i_s\}$ be the set of all indices k such that $\lambda_k > \lambda_{k+1}$. The i_j index precisely those rows of $Y(\lambda)$ where a box can be removed leaving a valid Young diagram of size $n - 1$. So let $\lambda^{(j)} = (\lambda_1, \dots, \lambda_{i_j} - 1, \dots, \lambda_r)$ ($1 \leq j \leq s$). Then each $\lambda^{(j)}$ is a partition of $n - 1$.

(8.9) Let $S^{n-1} \subseteq S^n$ be the stabilizer of 1, and let $\chi_{\lambda} \in \text{Irr}(S^n)$. With the above notation, $\chi_{\lambda}|_{S^{n-1}} = \sum_{j=1}^s \chi_{\lambda^{(j)}}$.

PROOF. For simplicity of notation, let X be the Specht module X_{λ} restricted to S^{n-1} , let $X_j = X_{\lambda^{(j)}}$, and put $\chi = \chi_{\lambda}|_{S^{n-1}}$ and $\chi_j = \chi_{\lambda^{(j)}}$. Observe that if T is a standard tableau of type λ and $T(i, j) = 1$, then $i = i_k$ for some

k and $j = \lambda_{i_k}$ because there cannot be a box directly below or to the right of the one occupied by 1. Deleting the box (i, j) (and subtracting one from each remaining entry) yields a tableau T' of type $\lambda^{(k)}$. Thus, we can define for each j a map $f_j : X \rightarrow X_j$ via

$$f_j(v(T)) = \begin{cases} v(T') & \text{if } T' \text{ is of type } \lambda^{(j)}, \\ 0 & \text{otherwise.} \end{cases}$$

If $\sigma \in S^{n-1}$ then it is immediate that $(T^\sigma)' = (T')^\sigma$ because σ fixes 1. It follows that $f_j \in \text{Hom}_{S^{n-1}}(X, X_j)$ and in particular that $(\chi, \chi_j) \neq 0$.

We now know that the character $\sum_{j=1}^s \chi_j$ is a constituent of χ . To complete the proof we only need to show that $\sum_{j=1}^s \chi_j(1) = \chi(1)$, but this is immediate from (8.7) since the map $T \rightarrow T'$ maps the standard tableaux of type λ bijectively to the disjoint union of the standard tableaux of type λ_j ($1 \leq j \leq s$). $\square$

(8.10) *Suppose that $n \geq 3$ and that λ and μ are partitions of n . Then there exists a partition π of $n-1$ such that $(\chi_\pi^S, \chi_\lambda) \neq (\chi_\pi^S, \chi_\mu)$.*

PROOF. Denote by $N(\lambda)$ the set of all Young diagrams which can be obtained by deleting one box from $Y(\lambda)$. By Frobenius reciprocity and (8.9) it suffices to show that for $n \geq 3$, $N(\lambda) = N(\mu)$ implies that $\lambda = \mu$.

Suppose that $\pi \in N(\lambda) \cap N(\mu)$ for some $\lambda \neq \mu$. Then there exist $i \neq j$ such that $\pi_i = \lambda_i = \mu_i - 1$, $\pi_j = \mu_j = \lambda_j - 1$, and $\pi_k = \lambda_k = \mu_k$ for $k \neq i, j$. Evidently these conditions characterize i and j , whence $|N(\lambda) \cap N(\mu)| \leq 1$. Thus, if $N(\lambda) = N(\mu)$ then $N(\lambda) = N(\mu) = \{\pi\}$. But this implies that $Y(\lambda)$ and $Y(\mu)$ are both rectangles, and it follows easily that $n = 2$. $\square$

CHAPTER 9

Symmetric Functions

In this section we make an apparent digression to develop the theory of symmetric functions. Not too surprisingly, however, this subject is intimately connected with the character theory of the symmetric group, as we shall see in §11. Since it would appear to be impossible to improve upon the superb exposition of [8], we will follow it closely.

Let Λ_n be the fixed subring of the action of S^n on $\mathbb{Z}[x_1, \dots, x_n]$ obtained by permuting the variables. We call Λ_n the *ring of symmetric polynomials in n variables*. Λ_n is a graded ring in the usual way:

$$\Lambda_n = \bigoplus_{k \geq 0} \Lambda_n^k$$

where Λ_n^k is the space of homogeneous symmetric polynomials of total degree k . It is clear that S^n permutes the monomials of degree k , and that each S^n -orbit contains a unique monomial $x^\lambda = \prod_{i=1}^n x_i^{\lambda_i}$, where $\lambda = (\lambda_1, \dots, \lambda_n)$ is a partition of k (with some parts possibly zero). We let $m_\lambda(x_1, \dots, x_n)$ be the sum of all the distinct conjugates of x^λ under S^n . Then the m_λ form a basis for Λ_n^k .

The partition λ can have no more than n parts, which is not a problem as long as $n \geq k$, but in order to remove this restriction in general we let the number of variables tend to infinity by defining $\Lambda^k = \varprojlim_n \Lambda_n^k$. More precisely, we let $\nu_n^k : \Lambda_{n+1}^k \rightarrow \Lambda_n^k$ be the map induced by specializing x_{n+1} to zero. Then Λ^k consists of all sequences $f = \{f_1, f_2, \dots\}$ such that $f_n \in \Lambda_n^k$ for all n and $f_n = \nu_n^k(f_{n+1})$. Hence, $f_{n+1} = f_n + (\text{monomials involving } x_{n+1})$ so we can think of the elements of $\Lambda^k = \Lambda^k(x)$ as formal infinite sums of monomials of total degree k in the variables $x = \{x_1, x_2, \dots\}$.

Note that we are taking a separate limit for each degree k , and as soon as $n \geq k$, ν_n^k is an isomorphism since a partition of k can have at most k parts. Hence Λ^k is a free $\mathbb{Z}$ -module whose rank is the number of partitions of k . To conserve notation we again denote by m_λ the unique element of Λ^k which projects onto $m_\lambda(x_1, \dots, x_n)$ for $n \geq k$. For example,

$$m_{(2,1)} = x_1^2 x_2 + x_2^2 x_1 + x_1^2 x_3 + x_3^2 x_1 + x_2^2 x_3 + x_3^2 x_2 + \dots$$

It is clear that the m_λ are a basis for Λ^k . Having defined Λ^k for all k , we now put

$$\Lambda = \Lambda(x) = \bigoplus_{k \geq 0} \Lambda^k.$$

There is an obvious way to define multiplication in Λ which makes the projection maps $\Lambda \rightarrow \Lambda_n$ ring homomorphisms. This product converts Λ to a graded ring which is called the *ring of symmetric functions* in the variables $x = \{x_1, x_2, \dots\}$.

In addition to the m_λ , we are going to introduce and study various other bases for Λ . These are, in their order of appearance:

- the elementary symmetric functions e_λ ,
- the complete symmetric functions h_λ ,
- the power sums p_λ ,
- the Schur functions s_λ .

We begin with the elementary symmetric functions by defining

$$E(t) = \prod_{i \geq 1} (1 + x_i t) = \sum_{r \geq 0} e_r t^r \in \Lambda[[t]];$$

then

$$e_r = \sum_{i_1 < i_2 < \dots < i_r} x_{i_1} x_{i_2} \cdots x_{i_r}.$$

Given a partition $\lambda = (\lambda_1, \dots, \lambda_s)$ we define $e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_s}$.

(9.1) *There exist nonnegative integers $a_{\lambda\mu}$ such that for any partition λ we have*

$$e_{\lambda'} = m_\lambda + \sum_{\mu < \lambda} a_{\lambda\mu} m_\mu,$$

where $<$ is the lexicographic order introduced just before (6.3). In particular, $\{e_\lambda : |\lambda| = k\}$ is a basis for Λ^k , the e_r are algebraically independent, and $\Lambda = \mathbb{Z}[e_1, e_2, \dots]$.

PROOF. Given the above formula, the stated consequences are immediate. To prove the formula, let c_i be the number of parts of λ' equal to i . If we expand the product $e_{\lambda'} = \prod_i e_i^{c_i}$ as a sum of monomials and order the monomials lexicographically with $x_1 > x_2 > \dots$, then the largest term is clearly

$$w = x_1^{c_1} \cdot (x_1 x_2)^{c_2} \cdots (x_1 x_2 \cdots x_r)^{c_r}$$

and it occurs with multiplicity 1. Since one can easily read off from the Young diagram that $c_i = \lambda_i - \lambda_{i+1}$, we have $\lambda_k = \sum_{i \geq k} c_i$ and hence $w = x^\lambda = \prod_i x_i^{\lambda_i}$. We conclude that when $e_{\lambda'}$ is expanded as a linear combination of the m_μ , the leading term is m_λ and it has multiplicity 1. $\square$

To obtain the complete symmetric functions, we set

$$H(t) = \prod_{i \geq 1} (1 - x_i t)^{-1} = \sum_{r \geq 0} h_r t^r.$$

Then $h_r \in \Lambda$, and using the expansion $(1 - x_i t)^{-1} = \sum_{k \geq 0} x_i^k t^k$, it follows easily that h_r is the sum of all monomials of total degree r . Given a partition $\lambda = (\lambda_1, \dots, \lambda_s)$, we define $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_s}$.

Notice that $E(t)H(-t) = 1$, which implies that

$$(9.2) \quad \sum_{r=0}^n (-1)^r h_r e_{n-r} = 0 \quad \text{for all } n \geq 1.$$

Since the e_r are algebraically independent, we can define a graded endomorphism $\omega : \Lambda \rightarrow \Lambda$ via $\omega(e_r) = h_r$. Applying ω to (9.2), setting $s = n - r$, and multiplying by -1 if n is odd, we obtain

$$\sum_{s=0}^n (-1)^s \omega(h_{n-s}) h_s = 0 \quad \text{for all } n \geq 1$$

which, together with (9.2) and an easy induction argument, implies that $\omega(h_r) = e_r$ for all r and thus $\omega^2 = 1$. In particular, we have

(9.3) *The complete symmetric functions $\{h_\lambda : |\lambda| = k\}$ are a basis for Λ^k , the h_r are algebraically independent, and $\Lambda = \mathbf{Z}[h_1, h_2, \dots]$. $\square$*

The r th power sum is defined by $p_r = \sum_{i \geq 1} x_i^r \in \Lambda$ for any $r \geq 1$, and the generating function is

$$P(t) = \sum_{r \geq 1} p_r t^{r-1}.$$

We see that $P(t)$ is the logarithmic derivative of $H(t)$:

$$\frac{H'(t)}{H(t)} = \sum_{i \geq 1} \frac{x_i}{1 - x_i t} = \sum_{i \geq 1} x_i \sum_{k \geq 0} x_i^k t^k = \sum_{k \geq 0} p_{k+1} t^k = P(t).$$

Taking the logarithmic derivative of the identity $E(t)H(-t) = 1$ we obtain

$$\frac{E'(t)}{E(t)} = \frac{H'(-t)}{H(-t)} = P(-t).$$

Extending the automorphism $\omega : \Lambda \rightarrow \Lambda$ to an automorphism of $\Lambda[[t]]$, we get $P^\omega(t) = P(-t)$, and it follows that

$$(9.4) \quad \omega(p_r) = (-1)^{r-1} p_r \quad \text{for } r \geq 1.$$

Given a partition $\lambda = (\lambda_1, \dots, \lambda_s)$ we define $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_s}$.

(9.5) $h_k = \sum_{|\lambda|=k} n_\lambda^{-1} p_\lambda$ for all k , where n_λ is defined in (6.6). In particular, the p_λ with $|\lambda| = k$ are a basis for $\Lambda^k \otimes \mathbf{Q}$.

PROOF. Let

$$f(t) = \exp \sum_{r \geq 1} \frac{p_r}{r} t^r.$$

Then $f'(t)/f(t) = P(t)$ and since $f(0) = H(0) = 1$, we have $f(t) = H(t)$. Thus,

$$H(t) = \exp \sum_{r \geq 1} \frac{p_r}{r} t^r = \prod_{r \geq 1} \exp \frac{p_r}{r} t^r = \prod_{r \geq 1} \sum_{n_r \geq 0} \frac{p_r^{n_r}}{n_r! r^{n_r}} t^{r n_r}.$$

Close inspection of the right-hand side of the above equation reveals that the coefficient of t^k is precisely $\sum_{|\lambda|=k} n_\lambda^{-1} p_\lambda$. In particular, it follows that $\mathbf{Q}[p_1, p_2, \dots] = \mathbf{Q}[h_1, h_2, \dots] = \Lambda \otimes \mathbf{Q}$, and thus the p_λ with $|\lambda| = k$ are a basis for $\Lambda^k \otimes \mathbf{Q}$. $\square$

CHAPTER 10

The Schur Functions

In this section, we continue our development of the theory of symmetric functions. We define the Schur functions, express them as alternating sums of complete symmetric functions, and use them as the orthonormal basis of a positive definite form on Λ . We first define a polynomial in n variables and then pass to the limit. Let $\alpha = (\alpha_1, \dots, \alpha_n)$ be any n -tuple of nonnegative integers, and let $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$ be the corresponding monomial, as before. Define

$$a_\alpha = \sum_{\sigma \in S^n} (-1)^\sigma x^{\sigma(\alpha)}.$$

Then a_α is *antisymmetric*, i.e., $\sigma(a_\alpha) = (-1)^\sigma a_\alpha$ for any $\sigma \in S^n$. Let $f = \sum_\alpha c_\alpha x^\alpha$ be an arbitrary antisymmetric polynomial. Then $c_{\sigma(\alpha)} = (-1)^\sigma c_\alpha$ and it follows that the a_α span the antisymmetric polynomials over $\mathbf{Z}$. Moreover, since f changes sign when the variables x_i and x_j are interchanged, f vanishes at the specialization $x_i = x_j$ and is therefore divisible by $x_i - x_j$ for all $i \neq j$. Thus, if we let $\delta = \delta_n = (n-1, n-2, \dots, 0)$, then f is divisible by the discriminant $a_\delta = \prod_{i < j} (x_i - x_j)$ and the quotient is a symmetric polynomial in n variables. Conversely, it is clear that if s is symmetric then $a_\delta s$ is antisymmetric, so multiplication by a_δ is a bijection from symmetric to antisymmetric polynomials. It follows that the symmetric polynomials a_α/a_δ span Λ_n over $\mathbf{Z}$.

If we assume that $\alpha_1 > \alpha_2 > \dots > \alpha_n$ in the definition of a_α , then a_α is of the form $a_{\lambda+\delta}$ where λ is a partition with at most n parts, and the symmetric polynomials $s_\lambda = a_{\lambda+\delta}/a_\delta$ span Λ_n over $\mathbf{Z}$. The polynomials s_λ are called *Schur polynomials*. By a dimension count we conclude that

(10.1) *The Schur polynomials are a $\mathbf{Z}$ -basis for Λ_n .*

We next want to express the Schur polynomials in terms of the complete symmetric polynomials. Let $\alpha = \lambda + \delta$ where, as above, λ is a partition with at most n parts. It is convenient to assume that λ has exactly n parts, by including additional zero terms if necessary. Let $\mathbf{A}_\alpha$ be the $n \times n$ matrix whose (i, j) entry is $x_i^{\alpha_j}$. Then $a_\alpha = \det \mathbf{A}_\alpha$.

Let $e_r^{(k)}$ be the r th elementary symmetric polynomial in the $n-1$ variables $\{x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n\}$ and put $E^{(k)}(t) = \sum_{r=0}^{n-1} e_r^{(k)} t^r$. Then

$$(10.2) \quad H_n(t) E^{(k)}(-t) = (1 - x_k t)^{-1}$$

where $H_n(t) = \prod_{i=1}^n (1 - x_i t)^{-1}$. Equating coefficients of t^{α_j} in (10.2) yields

$$(10.3) \quad x_k^{\alpha_j} = \sum_{i=1}^n (-1)^{n-i} h_{\alpha_j - n + i} e_{n-i}^{(k)}$$

with the convention (henceforth adopted) that $h_s = 0$ for $s < 0$. To rewrite (10.3) in matrix form, we let $\mathbf{E}$ be the $n \times n$ matrix whose (k, i) entry is $(-1)^{n-i} e_{n-i}^{(k)}$ and we let $\mathbf{H}_\alpha$ be the matrix whose (i, j) entry is $h_{\alpha_j - n + i}$. Then (10.3) becomes

$$\mathbf{A}_\alpha = \mathbf{E} \mathbf{H}_\alpha$$

and taking determinants we have $a_\alpha = \det \mathbf{E} \det \mathbf{H}_\alpha$. Note that if $\alpha_j < 0$ for some j , then the j th column of $\mathbf{H}_\alpha$ is zero, so we define $a_\alpha = 0$ if any $\alpha_j = 0$. Since $\mathbf{H}_\delta$ is lower triangular with $h_0 = 1$ on the diagonal, we obtain the critical formula

$$(10.4) \quad \frac{a_\alpha}{a_\delta} = \frac{\det \mathbf{E} \det \mathbf{H}_\alpha}{\det \mathbf{E} \det \mathbf{H}_\delta} = \det \mathbf{H}_\alpha = \sum_{\sigma \in S^n} (-1)^\sigma h_{\alpha - \sigma(\delta)}$$

for any n -tuple of integers α . If $\alpha = \lambda + \delta$ where λ is a partition, we get

$$(10.5) \quad s_\lambda = \frac{a_{\lambda+\delta}}{a_\delta} = \det \mathbf{H}_{\lambda+\delta} = \sum_{\sigma \in S^n} (-1)^\sigma h_{\lambda+\delta-\sigma(\delta)}.$$

For example (taking $n = 2$), we get $s_{(3,1)} = h_3 h_1 - h_4 h_0$, a result which may be checked directly from the definitions.

Notice that if λ has exactly r nonzero parts, then

$$\mathbf{H}_\alpha = \begin{bmatrix} \mathbf{H}_{11} & 0 \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix}$$

where $\mathbf{H}_{11}$ has dimension r and $\mathbf{H}_{22}$ is lower triangular with 1 on the diagonal. In particular, $\det \mathbf{H}_\alpha = \det \mathbf{H}_{11}$, so s_λ is a fixed polynomial of degree r in the h_i , independent of the number of variables. It follows that (10.5) defines in the limit a symmetric function s_λ which we call a *Schur function*.

If $|\lambda| = n$ then all the nonzero polynomials $h_{\lambda+\delta-\sigma(\delta)}$ are homogeneous of degree n . Moreover, $\lambda + \delta - \sigma(\delta) > \lambda$ for all $\sigma \neq 1$, and therefore (10.5) expresses s_λ as an “upper triangular sum” of h_μ with $|\mu| = |\lambda|$. Since the inverse of an upper triangular matrix is again upper triangular, we have

(10.6) *Let n be a nonnegative integer. Then for all partitions λ, μ with $|\lambda| = |\mu| = n$ there exist integers $b_{\lambda\mu}$ such that*

$$h_\lambda = s_\lambda + \sum_{\mu > \lambda} b_{\lambda\mu} s_\mu.$$

In particular, the Schur functions of degree n are a $\mathbb{Z}$ -basis for Λ^n .

Now recall that since Λ is a graded ring, it is embedded in its completion $\bar{\Lambda} = \prod_k \Lambda^k$ in just the same way that a polynomial ring is embedded in its ring of formal power series. Let $x = \{x_1, x_2, \dots\}$ and $y = \{y_1, y_2, \dots\}$ be two sets of indeterminates. We denote by $h_\lambda(x)$, $p_\lambda(y)$, etc. the various symmetric functions in the variables x and y , and by $H_x(t)$, $P_y(t)$, etc. their

respective generating functions. We are going to make some calculations in the ring $\overline{\Lambda}(x, y) = \overline{\Lambda}(x \cup y)$. Let

$$(10.7) \quad \begin{aligned} K(x, y) &= \prod_{i,j} (1 - x_i y_j)^{-1} \\ &= \prod_j H_x(y_j) = \prod_j \sum_r h_r(x) y_j^r = \sum_{\alpha} h_{\alpha}(x) y^{\alpha} \end{aligned}$$

where α ranges over all sequences of nonnegative integers with finite sum. If we denote by z the (countable) set of variables $\{x_i y_j\}$ then $K(x, y) = H_z(1)$, and $p_{\lambda}(z) = p_{\lambda}(x) p_{\lambda}(y)$. Hence, (9.5) yields

$$(10.8) \quad K(x, y) = H_z(1) = \sum_{\lambda} h_{\lambda}(z) = \sum_{\lambda} n_{\lambda}^{-1} p_{\lambda}(x) p_{\lambda}(y).$$

We also want to express $K(x, y)$ in terms of Schur functions. To do this, it seems that we must again work first with n variables and then pass to the limit. So specialize all but the first n variables to zero in (10.7) to get

$$K_n(x, y) = \prod_{i,j=1}^n (1 - x_i y_j)^{-1} = \sum_{\alpha} h_{\alpha}(x) y^{\alpha}$$

where α ranges over all n -tuples of integers (recall our conventions that $h_r = 0$ for $r < 0$, and $a_{\alpha} = 0$ if any $\alpha_j < 0$). Now multiply by $a_{\delta}(y) = \sum_{\sigma \in S^n} (-1)^{\sigma} y^{\sigma(\delta)}$ to get

$$(10.9) \quad a_{\delta}(y) K_n(x, y) = \sum_{\alpha, \sigma} (-1)^{\sigma} h_{\alpha}(x) y^{\alpha + \sigma(\delta)} = \sum_{\beta, \sigma} (-1)^{\sigma} h_{\beta - \sigma(\delta)}(x) y^{\beta}$$

where β ranges over all n -tuples of integers. From (10.4) we have

$$a_{\delta}(x) \sum_{\sigma} (-1)^{\sigma} h_{\beta - \sigma(\delta)} = a_{\beta}(x)$$

so multiplying (10.9) by $a_{\delta}(x)$ we obtain

$$a_{\delta}(x) a_{\delta}(y) K_n(x, y) = \sum_{\beta} a_{\beta}(x) y^{\beta}.$$

Since $a_{\beta}(x) = 0$ unless all entries of β are distinct and nonnegative, in which case $\beta = \sigma(\lambda)$ for some partition λ with at most n parts and some $\sigma \in S^n$, the right-hand sum can be rewritten to get

$$a_{\delta}(x) a_{\delta}(y) K_n(x, y) = \sum_{\lambda} a_{\lambda}(x) \sum_{\sigma} (-1)^{\sigma} y^{\sigma(\lambda)} = \sum_{\lambda} a_{\lambda}(x) a_{\lambda}(y)$$

where the sum is over partitions λ of at most n parts. Since $a_{\lambda} = 0$ unless λ has distinct parts, we get

$$K_n(x, y) = \sum_{\lambda} \frac{a_{\lambda+\delta}(x)a_{\lambda+\delta}(y)}{a_{\delta}(x)a_{\delta}(y)} = \sum_{\lambda} s_{\lambda}(x)s_{\lambda}(y)$$

summed over all partitions λ with at most n parts. Now letting $n \rightarrow \infty$ we have

$$(10.10) \quad K(x, y) = \sum_{\lambda} s_{\lambda}(x)s_{\lambda}(y).$$

Finally, we define a positive definite inner product $\langle \cdot, \cdot \rangle$ on Λ by requiring that Λ^j and Λ^k be orthogonal for $j \neq k$ and that the s_{λ} be an orthonormal basis for Λ^k . It is immediate from (10.6) that

$$(10.11) \quad \langle s_{\lambda}, h_{\lambda} \rangle = 1 \text{ and } \langle s_{\lambda}, h_{\mu} \rangle = 0 \text{ for } \mu < \lambda. \quad \square$$

The key fact we shall need subsequently is that the power sums are orthogonal. Let n_{λ} be the integer defined in (6.6). Then

$$(10.12) \quad \text{For any two partitions } \lambda, \mu \text{ of } n, \langle p_{\lambda}, p_{\mu} \rangle = \delta_{\lambda\mu} n_{\lambda}.$$

PROOF. Since the Schur functions of degree n are a $\mathbf{Z}$ -basis for Λ^n by (10.6), there are integers $c_{\lambda\nu}$ such that $p_{\lambda} = \sum_{\nu} c_{\lambda\nu} s_{\nu}$. Hence, $\langle p_{\lambda}, p_{\mu} \rangle = \sum_{\nu} c_{\lambda\nu} c_{\mu\nu}$. If we let $\mathbf{C}$ be the matrix whose (λ, μ) entry is $c_{\lambda\mu}$, then we want to show that $\mathbf{C}\mathbf{C}^t = \mathbf{N} = \text{diag}(n_{\lambda})$. From (10.8) and (10.10) we have

$$K(x, y) = \sum_{\lambda} s_{\lambda}(x)s_{\lambda}(y) = \sum_{\lambda} n_{\lambda}^{-1} p_{\lambda}(x)p_{\lambda}(y) = \sum_{\lambda, \nu, \mu} n_{\lambda}^{-1} c_{\lambda\nu} c_{\lambda\mu} s_{\nu}(x)s_{\mu}(y).$$

Put $d_{\nu\mu} = \sum_{\lambda} n_{\lambda}^{-1} c_{\lambda\nu} c_{\lambda\mu}$. Then

$$(10.13) \quad \sum_{\lambda} s_{\lambda}(x)s_{\lambda}(y) = \sum_{\nu, \mu} d_{\nu\mu} s_{\nu}(x)s_{\mu}(y).$$

Let $\bar{y}$ be a specialization of y in which all but a finite number of the y_j are zero. Then (10.13) becomes a dependence relation on the $s_{\lambda}(x)$, whence we get the relations

$$s_{\lambda}(\bar{y}) = \sum_{\mu} d_{\lambda\mu} s_{\mu}(\bar{y})$$

for all such specializations $\bar{y}$. But $s_{\lambda}(\bar{y})$ is just a specialization of a Schur polynomial in n variables, and since these are a basis for $\Lambda_n(y)$ by (10.1), it follows that $d_{\lambda\mu} = \delta_{\lambda\mu}$. In matrix form, this says $\mathbf{C}^t \mathbf{N}^{-1} \mathbf{C} = \mathbf{I}$. Since (9.5) implies that $\mathbf{C}$ is nonsingular the result follows. $\square$

Notice that by (9.4), (9.5), and (10.12) it is immediate that

$$(10.14) \quad \omega \text{ is an isometry.} \quad \square$$

CHAPTER 11

The Littlewood-Richardson Ring

In this section we define a graded ring whose homogeneous component of degree n is the $\mathbf{Z}$ -module of generalized characters of S^n (but the multiplication is *not* pointwise) and we prove that the ring so defined is isomorphic to the ring of symmetric functions. Under the isomorphism, the irreducible character χ_λ corresponds to the Schur function s_λ , and the permutation character ψ_λ corresponds to the complete symmetric function h_λ . Using this isomorphism, we obtain the so-called “determinant form” expressing χ_λ in closed form as an integral linear combination of the ψ_μ . From the determinant form we derive several formulas for character values.

Let L^n be the $\mathbf{Z}$ -module of generalized characters of S^n for $n \geq 1$, let L^0 be a one-dimensional space spanned by an element called 1, and let $L = \bigoplus_{n \geq 0} L^n$. We convert L to a graded ring as follows. Identify S^i and S^j with the subgroups of S^{i+j} which fix $\{i+1, \dots, i+j\}$ and $\{1, \dots, i\}$ respectively, then $S^i S^j = S^i \times S^j$. If $f \in L^i$ and $g \in L^j$, we let $f \# g$ be the class function on $S^i \times S^j$ defined by $f \# g(x, y) = f(x)g(y)$ for all $x \in S^i$ and $y \in S^j$. Note that if $f = f_1 + f_2$, then $f \# g = f_1 \# g + f_2 \# g$, whence $f \# g$ is a generalized character by (2.10).

We now define $fg = (f \# g)^{S^{i+j}}$. Then $(f, g) \mapsto fg$ is a bilinear map $L^i \times L^j \rightarrow L^{i+j}$ which converts L to a graded ring. The product is commutative because if we interchange the roles of i and j above, then $S^i \times S^j$ is conjugate to $S^j \times S^i$ in S^{i+j} . In order to show that the product is associative, we identify S^k with the subgroup of S^{i+j+k} fixing $\{1, \dots, i+j\}$ and let $h \in L^k$. We claim that

$$(fg) \# h = (f \# g \# h)^{S^{i+j} \times S^k}$$

as can be easily seen from (4.3) after observing that coset representatives for $S^i \times S^j$ in S^{i+j} are simultaneously coset representatives for $S^i \times S^j \times S^k$ in $S^{i+j} \times S^k$. It then follows that

$$(fg)h = (f \# g \# h)^{S^{i+j+k}} = f(gh).$$

We call L the *Littlewood-Richardson ring*. The structure constants $c_{\lambda\mu}^\nu$ for L defined by the equations

$$\chi_\lambda \chi_\mu = \sum_{\nu} c_{\lambda\mu}^\nu \chi_\nu$$

were studied by Littlewood and Richardson, who stated a famous combinatorial rule for evaluating them. For a nice treatment, see [8].

Each homogeneous component L^n of L has a natural inner product defined, and we extend these to an inner product on L by declaring L^i and L^j to be orthogonal for $i \neq j$.

We denote by $[n]$ the principal character $1_{S^n} \in L^n$. Then recalling that ψ_π is the permutation character induced from a Young subgroup of type π , it is immediate that

$$(11.1) \text{ If } \pi = (\pi_1, \pi_2, \dots, \pi_s) \text{ then } \psi_\pi = [\pi_1][\pi_2] \cdots [\pi_s]. \quad \square$$

Recall that to each element $\sigma \in S^n$ we have associated a partition $\bar{\sigma}$, the cycle type of σ . Let $\rho(\sigma) = p_{\bar{\sigma}} \in \Lambda$ be the power-sum associated to $\bar{\sigma}$. We now define the *characteristic map* $\mathbf{ch}: L \rightarrow \Lambda_{\mathbf{Q}} = \Lambda \otimes \mathbf{Q}$ as follows:

$$\mathbf{ch}(f) = (f, \rho) = \frac{1}{n!} \sum_{\sigma \in S^n} f(\sigma) \rho(\sigma) = \sum_{\pi} n_{\pi}^{-1} f(\sigma_{\pi}) p_{\pi} \quad \text{for } f \in L^n,$$

where for each partition π of n , σ_{π} is a representative element of S^n of type π , and n_{π} is the integer defined in (6.6). Since all irreducible characters of S^n are rational-valued, this definition makes sense. As the above formula indicates, $\mathbf{ch}(f)$ can be interpreted as the inner product of two $\Lambda_{\mathbf{Q}}$ -valued functions, since there is a natural embedding $\mathbf{Q} \rightarrow \Lambda_{\mathbf{Q}}$.

The main result of this section is

(11.2) *The map $\mathbf{ch}$ defines an isometric isomorphism of L onto Λ such that for each partition π , $\mathbf{ch}(\psi_{\pi}) = h_{\pi}$.*

PROOF. We first show that $\mathbf{ch}$ is an isometry. Let $f, g \in L^n$. Then using (10.12) and (6.6) we have

$$\begin{aligned} \langle \mathbf{ch}(f), \mathbf{ch}(g) \rangle &= \sum_{\pi, \rho} n_{\pi}^{-1} n_{\rho}^{-1} f(\sigma_{\pi}) g(\sigma_{\rho}) \langle p_{\pi}, p_{\rho} \rangle \\ &= \sum_{\pi} n_{\pi}^{-1} f(\sigma_{\pi}) g(\sigma_{\pi}) = (f, g). \end{aligned}$$

Next, we argue that $\mathbf{ch}$ is an injective ring homomorphism. The important point here is to observe that Frobenius reciprocity (3.1) is a formal calculation which holds equally well for $\Lambda_{\mathbf{Q}}$ -valued functions, so that if $f \in L^n$ and $g \in L^m$, we have

$$\mathbf{ch}(fg) = (f \# g^{S^{n+m}}, \rho) = (f \# g, \rho_{S^n \times S^m}) = \frac{1}{n!m!} \sum_{x, y} f(x) g(y) \rho(x, y).$$

Since it is clear that $\rho(x, y) = \rho(x)\rho(y)$ for $(x, y) \in S^n \times S^m$, we have $\mathbf{ch}(fg) = \mathbf{ch}(f)\mathbf{ch}(g)$. Since $\mathbf{ch}$ is a graded map, its kernel is also graded, but if $f \in L^n$ and $\mathbf{ch}(f) = 0$, then $(f, g) = 0$ for all $g \in L^n$ whence $f = 0$.

Finally, we see from (9.5) that

$$\mathbf{ch}([n]) = \sum_{\pi} n_{\pi}^{-1} p_{\pi} = h_n$$

and therefore $\mathbf{ch}(\psi_{\pi}) = h_{\pi}$ by (11.1). In particular, $\mathbf{ch}(L) = \Lambda$ by (7.2). $\square$

Now for each partition $\lambda = (\lambda_1, \dots, \lambda_r)$ we define a generalized character $[\lambda] \in L$ via the formula

$$(11.3) \quad [\lambda] = \det[\lambda_j + i - j]_{(1 \leq i, j \leq r)}.$$

By (10.5) and (11.2) we have $\mathbf{ch}([\lambda]) = \det(h_{\lambda_j + i - j}) = s_\lambda$. Since $\mathbf{ch}$ is an isometry, it follows that $\langle [\lambda], [\lambda] \rangle = \langle s_\lambda, s_\lambda \rangle = 1$, so $\pm[\lambda]$ is an irreducible character. By (10.11) and (11.2) we have $\langle [\lambda], \psi_\lambda \rangle = 1$ which implies that $[\lambda]$ (and not $-[\lambda]$) is an irreducible character, and moreover $\langle [\lambda], \psi_\mu \rangle = 0$ for $\mu < \lambda$. It now follows easily from (7.2) that $[\lambda] = \chi_\lambda$, whence

$$(11.4) \quad \chi_\lambda = [\lambda] \text{ and } \mathbf{ch}(\chi_\lambda) = s_\lambda \text{ for every partition } \lambda. \quad \square$$

This result expresses each irreducible character of S^n as an integral linear combination of permutation characters of Young subgroups in closed form. For example (keeping in mind the conventions $[0] = 1$ and $[k] = 0$ for $k < 0$) we have

$$\begin{aligned} \chi_{(3,2,1)} &= \det \begin{bmatrix} [3] & [1] & 0 \\ [4] & [2] & [0] \\ [5] & [3] & [1] \end{bmatrix} \\ &= [3]([2][1] - [3][0]) - [1]([4][1] - [5][0]) \\ &= [3][2][1] - [3][3] - [4][1][1] + [5][1]. \end{aligned}$$

We can use the determinant form to obtain results on the values of irreducible characters. For example, let σ be an n -cycle in S^n . Then σ is not conjugate to an element of any proper Young subgroup, whence $\psi_\lambda(\sigma) = 0$ unless $\lambda = (n)$. But since the nonzero entries in the matrix $([\lambda_j + i - j])$ are strictly increasing down the columns and strictly decreasing down the rows, it is easy to see that the unique largest one is $[\lambda_1 + r - 1]$. Moreover, since $\lambda_1 + r - 1 \leq n$ with equality iff $\lambda_2 = \lambda_3 = \dots = \lambda_r = 1$, there can be at most one term equal to $[n]$ in the expansion of any determinant of the form (11.3), and it occurs in the expansion of exactly one such determinant, namely

$$\det \begin{bmatrix} [n-r+1] & 1 & 0 & \cdots & 0 \\ [n-r+2] & [1] & 1 & \cdots & 0 \\ & & [1] & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ [n] & [r-2] & & \cdots & [1] \end{bmatrix}.$$

Expanding along the first column and evaluating at σ , we have

$$(11.5) \quad \text{Let } \sigma \in S^n \text{ be an } n\text{-cycle. Then}$$

$$\chi_\lambda(\sigma) = \begin{cases} (-1)^s & \text{if } \lambda = (n-s, 1^s), \\ 0 & \text{otherwise.} \end{cases} \quad \square$$

We can also use (11.3) and (11.4) to obtain a degree formula. We first observe that if $\lambda = (\lambda_1, \dots, \lambda_r)$ is a partition of n and if we define $\lambda! = \prod_{i=1}^r \lambda_i!$, then $\psi_\lambda(1) = |S : S_\lambda| = n!/\lambda!$. If we expand the determinant in (11.3) and evaluate

each term at 1, it is easy to see that the result is an integer determinant:

$$\chi_\lambda(1) = n! \det \frac{1}{(\mu_j - r + i)!} = \frac{n!}{\mu!} \det \frac{\mu_j!}{(\mu_j - r + i)!},$$

where $\mu_j = \lambda_j + r - j$. For example, when $r = 3$ we get

$$\det \begin{bmatrix} \mu_1(\mu_1 - 1) & \mu_2(\mu_2 - 1) & \mu_3(\mu_3 - 1) \\ \mu_1 & \mu_2 & \mu_3 \\ 1 & 1 & 1 \end{bmatrix} = \det \begin{bmatrix} \mu_1^2 & \mu_2^2 & \mu_3^2 \\ \mu_1 & \mu_2 & \mu_3 \\ 1 & 1 & 1 \end{bmatrix}$$

because the two matrices differ by an elementary row operation. In general, we need to evaluate a determinant of the form $\det(f_i(\mu_j))$ where f_i is a monic polynomial of degree i . Working upward from the last row, we can successively transform the rows of any such matrix to the rows of the Vandermonde matrix by elementary row operations of determinant 1. Thus, putting $\Delta(\mu) = \prod_{i < j} (\mu_i - \mu_j)$ we have

(11.6) *Let $\lambda = (\lambda_1, \dots, \lambda_r)$ be a partition of n and let $\mu = (\mu_1, \dots, \mu_r)$ where $\mu_j = \lambda_j + r - j$. Then*

$$\chi_\lambda(1) = \frac{n!}{\mu!} \Delta(\mu) = \frac{n!}{\prod_j \mu_j!} \prod_{i < j} (\mu_i - \mu_j). \quad \square$$

CHAPTER 12

Two Useful Formulas

In this section we derive two formulas from the determinant form for irreducible characters obtained in §11: the *hook-length formula* for the degrees, and the *Murnaghan-Nakayama formula* which can be used to compute values of individual irreducible characters without having to compute the entire character table, as in §7.

Given a partition λ we define, for each node (i, j) of the Young diagram $Y(\lambda)$, a subset $H_{ij}(\lambda)$ of $Y(\lambda)$ called the (i, j) -hook as follows:

$$H_{ij}(\lambda) = \{(i, k) : k \geq j\} \cup \{(k, j) : k \geq i\}.$$

We then set $h_{ij}(\lambda) = |H_{ij}(\lambda)| = \lambda_i + \lambda'_j - i - j + 1$. For example, in the diagrams below for $\lambda = (3, 3, 2, 1)$, we have marked $H_{21}(\lambda)$ and entered the value h_{ij} at each node:



12.1 (Frame, Robinson, and Thrall). *With the above notation,*

$$\chi_\lambda(1) = \frac{n!}{\prod_{i,j} h_{ij}(\lambda)}.$$

PROOF. Let $\lambda = (\lambda_1, \dots, \lambda_r)$. By (11.6) it suffices to show that

$$(12.2) \quad \frac{\mu!}{\Delta(\mu)} = \prod_{i,j} h_{ij}(\lambda),$$

where $\mu = \lambda + \delta$ and $\Delta(\mu) = \prod_{i < j} (\mu_i - \mu_j)$. Note that $\mu_i = \lambda_i + r - i = h_{i1}(\lambda)$. Arguing by induction, we treat two cases separately.

Case 1: $\lambda_r = 1$. We remove the last box in column 1:

Let $\lambda' = (\lambda_1, \dots, \lambda_{r-1})$. Then $\mu'_i = \mu_i - 1$ ($1 \leq i \leq r-1$) and

$$\Delta(\mu) = \prod_{1 \leq i < j \leq r} (\mu_i - \mu_j) = \Delta(\mu') \prod_i (\mu_i - \mu_r) = \Delta(\mu') \prod_i (\mu_i - 1) = \Delta(\mu') \prod_i \mu'_i.$$

Moreover, $h_{i1}(\lambda') = \mu'_i$ and $h_{ij}(\lambda') = h_{i,j}(\lambda)$ for $j > 1$.

Inductively we have

$$\frac{\mu'!}{\Delta(\mu')} = \prod_{i=1}^{r-1} \mu'_i \prod_{j>1} h_{ij}(\lambda).$$

Multiplying through by $\prod_{i=1}^{r-1} \mu_i / \mu'_i$, we have established (12.2) in Case 1.

Case 2: $\lambda_r > 1$. We remove column 1 completely:

Let $\lambda' = (\lambda_1 - 1, \dots, \lambda_r - 1)$. Then $\mu'_i = \mu_i - 1$ ($1 \leq i \leq r$), $\Delta(\mu) = \Delta(\mu')$, and $h_{ij}(\lambda') = h_{i,j+1}(\lambda)$ for all i, j . Inductively, we thus have

$$\frac{\mu'!}{\Delta(\mu')} = \frac{\mu'!}{\Delta(\mu)} = \prod_{i,j} h_{i,j+1}(\lambda).$$

Multiplying through by $\prod_i \mu_i = \prod_i h_{i1}(\lambda)$, we have established (12.2). $\square$

Next, given a partition $\lambda = (\lambda_1, \dots, \lambda_r)$ and integers k, m with $1 \leq k \leq r$, we let $\lambda^{(k,m)}$ be the r -tuple $(\lambda_1, \dots, \lambda_{k-1}, \lambda_k - m, \dots, \lambda_r)$ and we define $\psi_{\lambda^{(k,m)}} = [\lambda_1] \cdots [\lambda_{k-1}] [\lambda_k - m] \cdots [\lambda_r]$. Note that $\psi_{\lambda^{(k,m)}} = 0$ unless $\lambda_k \geq m$.

(12.3) Suppose that σ is an m -cycle and π is disjoint from σ . If λ has r parts, then

$$\psi_{\lambda}(\pi\sigma) = \sum_{k=1}^r \psi_{\lambda^{(k,m)}}(\pi).$$

PROOF. Let $\mu = (\mu_1, \dots, \mu_s)$ be the type of $\pi\sigma$. Then $\mu_i = m$ for some i and π is of type $(\mu_1, \dots, \mu_{i-1}, \mu_{i+1}, \dots, \mu_s)$. Referring to (7.3), we only need observe that the set R of all maps $f: \Omega^s \rightarrow \Omega^r$ such that f is a μ -refinement of λ is the disjoint union of the sets $R^{(k)} = \{f \in R \mid f(i) = k\}$. $\square$

Given any sequence $\lambda = (\lambda_1, \dots, \lambda_r)$ of integers, we define the $r \times r$ matrix

$$D_{\lambda} = [\lambda_j + i - j]_{1 \leq i, j \leq r}$$

with entries in L , and we set $[\lambda] = \det(D_{\lambda}) \in L$. It is easy to check that $[\lambda]$ is homogeneous of degree $n = \sum_{j=1}^r \lambda_j$ and is therefore a generalized character of S^n . If λ is a partition, then $[\lambda] = \chi_{\lambda}$ by (11.4), and we will call the matrix D_{λ} standard.

(12.4) Continue the notation of (12.3). Then

$$\chi_{\lambda}(\pi\rho) = \sum_{k=1}^r [\lambda^{(k,m)}](\pi).$$

PROOF. Expand the determinants, sum corresponding terms, and use (12.3). $\square$

To obtain the Murnaghan-Nakayama formula from (12.4), we need to understand the generalized character $[\lambda]$ when λ is not a partition. Putting $\mu_j = \lambda_j + r - j$ as usual, we notice that the bottom row of D_{λ} is $([\mu_1], \dots, [\mu_r])$. Denote by $\langle a \rangle$ the column vector of length r whose j th entry is $[a - r + j]$ for any integer $\langle a \rangle$. We will call any such column uniform. In particular, the

columns of D_λ are uniform, and

$$D_\lambda = (\langle \mu_1 \rangle, \langle \mu_2 \rangle, \dots, \langle \mu_r \rangle).$$

Conversely, any matrix $D' = (\langle \mu'_1 \rangle, \dots, \langle \mu'_r \rangle)$ with uniform columns is of the form $D_{\lambda'}$ where $\lambda' = \text{diag}(D')$. Moreover, $\mu'_j - \mu'_{j+1} = \lambda'_j - \lambda'_{j+1} + 1$, so λ' is a partition iff the μ'_j are strictly decreasing. It follows that if the μ'_j are distinct, then the columns of D' are some permutation τ of the columns of some standard matrix D_λ and hence $[\lambda'] = (-1)^\tau \chi_\lambda$. If the μ'_j are not distinct, then of course $[\lambda'] = 0$.

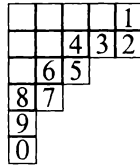
Returning now to the generalized characters in (12.4), consider the matrix $D_{\lambda^{(k,m)}} = (\langle \mu_1 \rangle, \dots, \langle \mu_k - m \rangle, \dots, \langle \mu_r \rangle)$. The μ_j are strictly decreasing, and we may as well assume that there is some index $i \geq k$ such that $\mu_i > \mu_k - m > \mu_{i+1}$, otherwise $[\lambda^{(k,m)}] = 0$. The column permutation required to bring $D_{\lambda^{(k,m)}}$ to standard form is then $(i, i-1, \dots, k)$ and if we let $\lambda^{(k)}$ be the diagonal of the resulting standard matrix, then

$$(12.5) \quad [\lambda^{(k,m)}] = (-1)^{i-k} \chi_{\lambda^{(k)}}.$$

To obtain $D_{\lambda^{(k)}}$ from $D_{\lambda^{(k,m)}}$, columns $k+1$ through i are shifted left one position, and column k is shifted right $i-k$ positions. It follows that

$$\lambda_j^{(k)} = \lambda_{j+1} - 1 \quad \text{for } k \leq j < i, \quad \text{and} \quad \lambda_i^{(k)} = \lambda_k - m + i - k.$$

The above formula is a prescription for converting a partition λ of n to a partition $\lambda^{(k)}$ of $n-m$. In terms of removing nodes from the Young diagram, it says to remove $\lambda_j - \lambda_{j+1} + 1$ nodes from row j for $k \leq j < i$, and if the total number of nodes thus removed is t , to remove a further $m-t$ nodes from row i . For example, in the diagram below with $\lambda = (5^2, 3, 2, 1^2)$ and $m = 6$, $\lambda^{(2)} = (5, 2, 1^3)$ is obtained by removing boxes 2 through 7, $\lambda^{(3)} = (5^2)$ is obtained by removing boxes 5 through 0, and $\lambda^{(k)} = 0$ for all other k .



Evidently, we obtain $Y(\lambda^{(k)})$ by removing a connected segment of the “rim” of $Y(\lambda)$ of length m beginning in row k if the resulting diagram is valid, otherwise $\lambda^{(k)} = 0$. Such a connected rim segment is called a *skew m -hook* of λ . Given a partition λ we will say that a skew hook s is *removable* from λ if the resulting Young diagram is valid, and we will denote the resulting partition by $\lambda \setminus s$. We let $l(s)$ be the number of rows spanned by s . Combining (12.4) and (12.5) with the above discussion, we have

12.6 (Murnaghan-Nakayama). *Suppose that ρ is an m -cycle and π is disjoint from ρ . Then*

$$\chi_\lambda(\pi\rho) = \sum_s (-1)^{l(s)-1} \chi_{\lambda \setminus s}(\pi)$$

where the sum ranges over all skew m -hooks s which are removable from λ . $\square$

The Murnaghan-Nakayama formula is a generalization of (8.9), which is the special case $m = 1$. The formula can be used to recursively compute $\chi_\lambda(\sigma)$ for any partition λ , and any permutation σ . For example, if $\lambda = (5^2, 3, 2, 1^2)$ as above and σ is of type $(6, 5, 5, 1)$, write $\sigma = \sigma_1 \sigma_2 \sigma_3$ where the σ_i are disjoint cycles of types 6, 5, 5, respectively. Since λ has two removable skew 6-hooks,

$$\chi_\lambda(\sigma) = \chi_{(5, 2, 1^4)}(\sigma_2 \sigma_3) - \chi_{(5^2, 1)}(\sigma_2 \sigma_3).$$

Continuing, $(5, 2, 1^4)$ has one removable skew 5-hook, which gives

$$\chi_{(5, 2, 1^4)}(\sigma_2 \sigma_3) = -\chi_{(1^6)}(\sigma_3) = 1$$

because σ_3 is even, while $(5^2, 1)$ also has one removable skew 5-hook, which gives

$$\chi_{(5^2, 1)}(\sigma_2 \sigma_3) = -\chi_{(4, 1^2)}(\sigma_3).$$

Since $(4, 1^2)$ has no removable skew 5-hooks, $\chi_{(4, 1^2)}(\sigma_3) = 0$. We conclude that $\chi_\lambda(\sigma) = 1$.

Part III

CHAPTER 13

The Hecke Algebra

In this section we define the Hecke algebra¹ (of type A_{n-1}) and prove that it is isomorphic to the group algebra $\mathbf{Q}[t]S^n$. We begin by summarizing some basic facts that we shall need about S^n .

(13.1) *Let $\sigma_i = (i, i+1) \in S^n$ for $1 \leq i \leq n$. Then*

$$(i) \quad \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \text{ for } 1 \leq i < n,$$

$$(ii) \quad \sigma_i \sigma_j = \sigma_j \sigma_i \text{ for } |i - j| \geq 2,$$

$$(iii) \quad \sigma_i^2 = 1 \text{ for } 1 \leq i < n.$$

Moreover, the above relations are a presentation for S^n .

PROOF. The relations are easily verified. The fact that they are a presentation is proved in [2]. $\square$

We now define an algebra $H_n = H_n[t]$ over the polynomial ring $\mathbf{Q}[t]$ with generators $g_1, g_2, \dots, g_{n-1}$ subject to the relations

$$(i) \quad g_i g_{i+1} g_i = g_{i+1} g_i g_{i+1} \text{ for } 1 \leq i < n,$$

$$(ii) \quad g_i g_j = g_j g_i \text{ for } |i - j| \geq 2,$$

$$(iii) \quad g_i^2 = (t-1)g_i + t \text{ for } 1 \leq i < n.$$

Given a complex number q , we denote by $H_n[q]$ the specialization obtained by setting $t = q$. It is clear from (13.1) that $H_n[1] = \mathbf{Q}S^n$. We set $R = \mathbf{Q}(t)$, the field of rational functions, and $RH_n = H_n \otimes R$. For convenience, we write

$$g_{i,j} = \begin{cases} g_{n-i} g_{n-i-1} \cdots g_{n-j} & \text{if } i \leq j, \\ 1 & \text{if } i > j. \end{cases}$$

(13.2) *H_n is spanned over $\mathbf{Q}[t]$ by elements of the form $w g_{1,i}$ for some $i \geq 0$, where $w \in \langle g_1, \dots, g_{n-2} \rangle$.*

PROOF. We first argue that an element of the form $g_{n-1} w g_{n-1}$ with $w \in \langle g_1, \dots, g_{n-2} \rangle$ can be rewritten as a $\mathbf{Q}[t]$ -linear combination of words involv-

¹The Hecke algebra receives its name from the following fact, which we shall not prove here. Let $G = \mathbf{GL}(n, q)$, the general linear group over the finite field of q elements. Let $B \subseteq G$ be the upper-triangular subgroup, and for each $x \in G$ let $\hat{x}$ denote the sum in $\mathbf{Q}G$ of the elements of the double coset BxB . Then the subalgebra generated by the $\hat{x}$ is isomorphic to $H_n[q]$. See [2] for details.

ing at most one occurrence of g_{n-1} . Namely, if $w \in \langle g_1, \dots, g_{n-3} \rangle$ then

$$g_{n-1}w g_{n-1} = w g_{n-1}^2 = (t-1)w g_{n-1} + tw.$$

Otherwise, we may assume by induction on n that $w = w_1 g_{n-2} w_2$ where $w_i \in \langle g_1, \dots, g_{n-3} \rangle$ ($i = 1, 2$), and then

$$\begin{aligned} g_{n-1}w g_{n-1} &= g_{n-1}w_1 g_{n-2} w_2 g_{n-1} = w_1 g_{n-1} g_{n-2} g_{n-1} w_2 \\ &= w_1 g_{n-2} g_{n-1} g_{n-2} w_2. \end{aligned}$$

It follows that H_n is spanned by $\langle g_1, \dots, g_{n-2} \rangle$ together with words of the form $w_1 g_{n-1} w_2$ with $w_i \in \langle g_1, \dots, g_{n-2} \rangle$. By induction on n , w_2 is a $\mathbf{Q}[t]$ -linear combination of words of the form $w_3 g_{2,i}$ for some $i \geq 1$ where $w_3 \in \langle g_1, \dots, g_{n-3} \rangle$, and since $w_3 g_{n-1} = g_{n-1} w_3$ the result follows. $\square$

Inductively, we define $w \in H_n$ to be a *standard word* if it is of the form $w_1 g_{1,i}$ for some $i \geq 0$, where w_1 is a standard word in $\langle g_1, \dots, g_{n-2} \rangle$.

(13.3) *For each $\sigma \in S^n$ there is a unique standard word $w_\sigma \in H_n$ with $w_\sigma[1] = \sigma$, and the standard words are a free $\mathbf{Q}[t]$ -basis for H_n . Furthermore, $H_n \otimes_{\mathbf{Q}[t]} F$ is semisimple for any field $F \supseteq \mathbf{Q}[t]$.*

PROOF. We use induction on n . Let $\hat{H}_{n-1} = \langle g_1, \dots, g_{n-2} \rangle \subseteq H_n$. Then $\hat{H}_{n-1}$ is a homomorphic image of H_{n-1} and is therefore certainly spanned by the standard words in $g_1, \dots, g_{n-2}$. By (13.2) the standard words span H_n . If we let $x[1]$ be the image of x in $H_n[1]$, then $w[1] \in S^n$ for any word w in the g_i , but since the standard words span H_n their images must span $\mathbf{Q}S^n$ and it follows that

$$S^n = \{w[1] : w \text{ is a standard word in } H_n\}.$$

On the other hand, an immediate induction argument shows that there are just $n!$ standard words, so for $\sigma \in S^n$ we can define w_σ to be the unique standard word such that $w_\sigma[1] = \sigma$. Moreover, if there were a relation

$$\sum_{\sigma} p_{\sigma}(t) w_{\sigma} = 0$$

with $p_{\sigma}(t) \in \mathbf{Q}[t]$ and $\gcd\{p_{\sigma}(t)\} = 1$, we could set $t = 0$ and deduce that each $p_{\sigma}(t)$ is divisible by t , which is impossible.

Let $\Delta_n(t)$ be the determinant of the trace form with respect to the basis of standard words. Then $\Delta_n(1)$ is the discriminant of $\mathbf{Q}S^n$ which is non-zero, whence $\Delta_n(t) \neq 0$. Since $\{w_{\sigma} \otimes 1\}$ is an F -basis for $H_n \otimes_{\mathbf{Q}[t]} F$ for any extension field F of $\mathbf{Q}[t]$, it follows that $H_n \otimes_{\mathbf{Q}[t]} F$ is semisimple. $\square$

In particular, (13.3) implies that the natural maps $H_n \rightarrow RH_n$ and $H_{n-1} \rightarrow H_n$ are embeddings. Henceforth, we shall make the necessary identifications to make these maps inclusions. We will write $H = H_n$ when there is no danger of confusion.

Now let $P = \mathbf{Q}[[t-1]]$ be the ring of formal power series in $(t-1)$ and let $L = \mathbf{Q}((t-1))$ be its field of fractions, the field of formal Laurent series in $(t-1)$. We denote by PH and LH the extended algebras $H \otimes P$ and $H \otimes L$,

respectively. Given a module M for H or PH , we denote by $M[1]$ the S^n -module obtained by specializing at $t = 1$. More precisely, we “extend scalars” by the ring homomorphism induced by setting $t = 1$ to obtain $M[1] = M \otimes \mathbf{Q}S^n$.

(13.4) *Let π be a partition of n . Then there exists a P -free PH -module M_π such that $M_\pi[1]$ is an irreducible $\mathbf{Q}S^n$ -module affording the character χ_π . Moreover, the LH -modules $\{M_\pi \otimes L : \pi \text{ a partition of } n\}$ are a complete set of pairwise nonisomorphic absolutely irreducible LH -modules.*

PROOF. This result is proved using the important technique of *lifting idempotents*. Let $I_\pi \subseteq \mathbf{Q}S^n$ be a minimal right ideal affording χ_π . Then $\mathbf{Q}S^n = I_\pi \oplus J_\pi$ for some right ideal J_π by (1.6). Let $1 = e + f$ with $e \in I_\pi$ and $f \in J_\pi$. Then for any $x \in I_\pi$ we have $x = 1x = ex + fx$. Hence $fx = x - ex \in I_\pi \cap J_\pi = 0$ and $x = ex$. In particular, $e^2 = e$ and $I_\pi = e\mathbf{Q}S^n$. Choose an element $e_1 \in H$ with $e_1[1] = e$. Then $e_1^2 \equiv e_1 \pmod{(t-1)H}$. Inductively, we will construct a sequence $\{e_i\}$ of elements of H such that

- (i) $e_i^2 \equiv e_i \pmod{(t-1)^i H}$ for $i \geq 1$,
- (ii) $e_i \equiv e_{i-1} \pmod{(t-1)^{i-1} H}$ for $i \geq 2$.

Assuming that $\{e_1, \dots, e_i\}$ has already been constructed, let $y = e_i^2 - e_i$ and define $e_{i+1} = e_i + (1 - 2e_i)y$. Since $y \in (t-1)^i H$, e_{i+1} satisfies (ii). Moreover, since y commutes with e_i we have

$$\begin{aligned} e_{i+1}^2 - e_{i+1} &= e_i^2 - e_i + 2e_i(1 - 2e_i)y - (1 - 2e_i)y + (1 - 2e_i)^2 y^2 \\ &\equiv y + (2e_i - 1)(1 - 2e_i)y \pmod{(t-1)^{2i} H} \\ &= (1 - (1 - 2e_i)^2)y = -4y^2 \equiv 0 \pmod{(t-1)^{i+1} H}, \end{aligned}$$

thus completing the induction.

Condition (ii) now implies that there is an element $e_\infty \in PH$ such that $e_\infty \equiv e_i \pmod{(t-1)^i PH}$. Then (i) implies that $e_\infty^2 = e_\infty$. Let $M_\pi = e_\infty PH$. Then since $e_\infty[1] = e$ it follows that $M_\pi[1] = I_\pi$.

Let $J_1 = (1 - e_\infty)PH$. Then $PH = M_\pi \oplus J_1$ and $J_1[1] = J_\pi$. If f is another idempotent which generates a minimal right ideal of $\mathbf{Q}S^n$ contained in J_π , then the above technique will lift it to an idempotent f_∞ of J_1 . Continuing in this way, it is clear that any direct sum decomposition of $\mathbf{Q}S^n$ lifts to a direct sum decomposition of PH . Since $\mathbf{Q}$ is a splitting field for S^n by the results of §8, (1.7) and the above imply that there is a direct sum decomposition

$$(13.5) \quad PH = \bigoplus_{\pi} \widehat{M}_\pi$$

where $\widehat{M}_\pi$ is the direct sum of $\chi_\pi(1)$ copies of M_π .

Since P is a P.I.D. and M_π is a submodule of the free P -module PH , M_π is P -free. Tensoring (13.5) with L , we obtain a direct sum decomposition of LH in which $M_\pi \otimes L$ occurs with multiplicity at least equal to its degree. By (1.8) the proof is complete. $\square$

We next argue that R , the field of rational functions, is a splitting field for H , and that the modules M_π are actually writable over $\mathbf{Q}[t]$. This result

is really not necessary in the sequel, and is included basically for the sake of completeness.

(13.6) *For each partition π of n there exists a $\mathbf{Q}[t]$ -free H_n -module X_π such that $X_\pi[1]$ affords χ_π . Moreover, $\{X_\pi \otimes R : \pi \text{ a partition of } n\}$ is a complete set of pairwise nonisomorphic absolutely irreducible RH -modules.*

PROOF. We proceed by induction on n , the case $n = 2$ being trivial. For each partition π of n , let $N_\pi = M_\pi \otimes L$ where M_π is the PH -module constructed in (13.4). From (13.3) we see that H is a free left H_{n-1} -module of rank n on the basis $\{g_{1,i}\}$, which easily implies that if X is an H_{n-1} -module, then the induced module $X^H = X \otimes_{H_{n-1}} H$ has the property

$$X^H[1] = X[1]^S.$$

Then using (8.10) and induction it follows that for each pair of partitions $\lambda \neq \mu$ of n , there exists a partition π of $n-1$ and an H -module $X_{\lambda,\mu} = X_\pi^H$ with the following property:

$X_{\lambda,\mu} \otimes L$ contains N_λ and N_μ with different multiplicity.

Indeed, it is clear from (8.9) that one of these multiplicities is one and the other is zero. By (1.7) there is for each λ a unique irreducible RH -module Y_λ such that $Y_\lambda \otimes L$ has N_λ as a constituent. Assume, by way of contradiction, that $Y_\lambda = Y_\mu$ for some $\mu \neq \lambda$. Then Y_λ must be a constituent of $X_{\lambda,\mu} \otimes R$. But then $Y_\lambda \otimes L = Y_\mu \otimes L$ has only one of $\{N_\lambda, N_\mu\}$ as a constituent, which is impossible. It follows that $Y_\lambda \otimes L$ must be a multiple of N_λ for all λ .

Now using (8.9) and induction, there is an H -module $X_\lambda = X_\pi^H$ for suitable π such that $X_\lambda \otimes L$ contains N_λ with multiplicity one. This implies that $Y_\lambda \otimes L = N_\lambda$, and proves that R is a splitting field for H_n .

We next argue that $Y_\lambda \cong X_\lambda \otimes R$ for some $\mathbf{Q}[t]$ -free H -module X_λ . Namely, choose a nonzero element $x_1 \in Y_\lambda$, and let X_λ be the $\mathbf{Q}[t]$ -submodule spanned by the elements $\{x_1 w_\sigma : \sigma \in S^n\}$. Then X_λ is H -invariant, and it is also finitely generated and torsion-free over $\mathbf{Q}[t]$, so X_λ is $\mathbf{Q}[t]$ -free. Since $X_\lambda \subseteq Y_\lambda$, there is a natural map $\varphi : X_\lambda \otimes R \rightarrow Y_\lambda$, which is surjective since Y_λ is irreducible. Since R is the field of quotients of $\mathbf{Q}[t]$, it follows that for each $y \in Y_\lambda$ there exists $p \in \mathbf{Q}[t]$ such that $py \in X_\lambda$. This implies that a $\mathbf{Q}[t]$ -basis for X_λ remains linearly independent over R , whence $\text{rank}(X_\lambda) = \dim(Y_\lambda)$ and thus φ is an isomorphism.

Finally, we argue that $X_\pi[1] \cong M_\pi[1]$. Fix a partition π and let $M = M_\pi$ and $M_1 = X_\pi \otimes P$. It suffices to show that $M[1] = M_1[1]$. Since $N_\pi = M_\pi \otimes L \cong X_\pi \otimes L$, we can identify M and M_1 with cotorsion P -submodules of N_π . Then $M/(M \cap M_1) \cong (M + M_1)/M_1$ is a finitely generated P -torsion module and is therefore annihilated by some $p \in P$. By a careful choice of p , we may assume that $pM \subseteq M_1$ and $pM \not\subseteq (t-1)M_1$. Since $M_1/(t-1)M_1$ is irreducible, $M_1 = pM + (t-1)M_1$, whence

$$M/(t-1)M \cong pM/p(t-1)M \cong M_1/(t-1)M_1. \quad \square$$

Now let ξ_π be the character of H afforded by X_π . Note that $\xi_\lambda(w_\sigma)[1] = \chi_\lambda(\sigma)$ for all $\sigma \in S^n$. We are interested in computing the restriction of ξ_λ to the subalgebra $H_{n,k}$ of H_n generated by all the g_i except for g_k .

(13.7) Let $c_{\lambda\mu}^\pi$ be the structure constants of the Littlewood-Richardson ring with respect to the basis of irreducible characters. If $\sigma \in S^n$ fixes $\{1, \dots, k\}$ and $\rho \in S^n$ fixes $\{k+1, \dots, n\}$ then

- (i) $w_{\sigma\rho} = w_\sigma w_\rho$,
- (ii) the set of all such $w_{\sigma\rho}$ is a basis for $H_{n,k} \cong H_k \otimes H_{n-k}$, and
- (iii)

$$\xi_\pi(w_{\sigma\rho}) = \sum_{\lambda, \mu} c_{\lambda\mu}^\pi \xi_\lambda(w_\sigma) \xi_\mu(w_\rho) \quad \text{for all partitions } \pi \text{ of } n,$$

where the sum ranges over all partitions λ of k and μ of $n-k$.

PROOF. Since $\langle g_1, \dots, g_{k-1} \rangle \cong H_k$, and $\langle g_{k+1}, \dots, g_{n-1} \rangle \cong H_{n-k}$, and these subalgebras commute elementwise, there is an obvious epimorphism $\varphi : H_k \otimes H_{n-k} \rightarrow H_{n,k}$. On the other hand, it is clear from the definition of standard words that if w_1 is a standard word in $\{g_1, \dots, g_{k-1}\}$ and w_2 is a standard word in $\{g_{k+1}, \dots, g_{n-1}\}$ then w_1, w_2 , and $w_1 w_2$ are all standard words in $\{g_1, \dots, g_{n-1}\}$ and hence linearly independent by (13.3). It follows that φ is an isomorphism, and then that $\{w_\sigma : \sigma \in S^k \times S^{n-k}\}$ is a $\mathbf{Q}[t]$ -basis for $H_{n,k}$. Moreover, if we identify S^k and S^{n-k} with the obvious subgroups of S^n then $w_{\sigma\rho} = w_\sigma w_\rho$ for $\sigma \in S^k$ and $\rho \in S^{n-k}$.

By (1.9) $RH_{n,k}$ is semisimple and its irreducible modules are of the form $X_\lambda \otimes X_\mu$ where λ is a partition of k and μ is a partition of $n-k$. Denote by $\xi_{\lambda\mu}$ the character afforded by $X_\lambda \otimes X_\mu$. If $x \in H_k$ and $y \in H_{n-k}$, then $\xi_{\lambda\mu}(xy) = \xi_\lambda(x) \xi_\mu(y)$. Moreover, there exist nonnegative integers $b_{\lambda\mu}^\pi$ such that

$$\xi_\pi|_{H_{n,k}} = \sum_{\lambda, \mu} b_{\lambda\mu}^\pi \xi_{\lambda\mu}.$$

Let $\sigma \in S^k$ and $\rho \in S^{n-k}$. Then since $w_\sigma w_\rho = w_{\sigma\rho}$, we have

$$\xi_\pi(w_{\sigma\rho}) = \sum_{\lambda, \mu} b_{\lambda\mu}^\pi \xi_{\lambda\mu}(w_{\sigma\rho}) = \sum_{\lambda, \mu} b_{\lambda\mu}^\pi \xi_\lambda(w_\sigma) \xi_\mu(w_\rho).$$

To identify the $b_{\lambda\mu}^\pi$ as the Littlewood-Richardson constants $c_{\lambda\mu}^\pi$, we just specialize at $t = 1$:

$$\chi_\pi(\sigma\rho) = \sum_{\lambda, \mu} c_{\lambda\mu}^\pi \chi_\lambda(\sigma) \chi_\mu(\rho),$$

and use Frobenius reciprocity. $\square$

CHAPTER 14

The Markov Trace

In this section we define a $\mathbf{Q}[t]$ -linear function $\tau: H_n \rightarrow \mathbf{Q}[s, t]$ where s is another indeterminate, and we show how to express τ as a $\mathbf{Q}[s, t]$ -linear combination of characters of H_n . These results are originally due to Ocneanu [3]; however, our treatment follows [9].

(14.1) *There is a unique $\mathbf{Q}[t]$ -linear function $\tau_n: H_n \rightarrow \mathbf{Q}[s, t]$ with the following properties:*

- (i) $\tau_n(1) = 1$,
- (ii) $\tau_n(xy) = \tau_n(yx)$,
- (iii) $\tau_n(xg_{n-1}) = s\tau_{n-1}(x)$ for $x \in H_{n-1}$.

PROOF. Inductively, we may assume that there is a unique such function τ_{n-1} defined on H_{n-1} . Then (13.3) shows that there is at most one extension to H_n satisfying (ii) and (iii) above, namely we define

$$\tau_n(wg_{1,i}) = s\tau_{n-1}(wg_{2,i})$$

where w is a standard word in H_{n-1} , and then extend by $\mathbf{Q}[t]$ -linearity. Then $\tau_n|_{H_{n-1}} = \tau_{n-1}$ and we will write $\tau = \tau_n$ without danger of confusion. It is clear that τ satisfies (iii). In fact, we claim that τ satisfies

$$(14.2) \quad \tau(xg_{n-1}y) = s\tau(xy) \quad \text{for all } x, y \in H_{n-1}.$$

Namely, by linearity it suffices to show this when y is a standard word of the form $wg_{2,i}$ where $w \in H_{n-2}$. Then $xg_{n-1}y = xwg_{1,i}$ and (14.2) follows by writing xw as a $\mathbf{Q}[t]$ -linear combination of standard words and applying the definition of τ .

The remaining problem is to verify (ii). It suffices to show that

$$(14.3) \quad \tau(wg_i) = \tau(g_iw)$$

for all standard words w and all i . Suppose first that $w \in H_{n-1}$. If $i < n-1$ then (14.3) holds by induction. Let $w = w_1g_{2,j}$ where $w_1 \in H_{n-2}$. Then $g_{n-1}w = w_1g_{1,j}$ is a standard word and so is wg_{n-1} , so (14.3) holds by definition in this case.

So we may assume that $w = w_1g_{1,j}$ in (14.3) for some standard word w_1 in H_{n-1} and some $j \geq 2$. If $i < n-1$ then (14.3) follows by first applying (14.2) and then using (ii) in H_{n-1} . Thus, we are reduced to proving

$$(14.4) \quad \tau(g_{n-1}w_1g_{1,j}) = \tau(w_1g_{1,j}g_{n-1}).$$

Let $w_1 = w_2 g_{2,k}$ where $w_2 \in H_{n-2}$. Then

$$\begin{aligned}
 \tau(g_{n-1} w_1 g_{1,j}) &= \tau(g_{n-1} w_2 g_{2,k} g_{1,j}) \\
 &= \tau(w_2 g_{n-1} g_{n-2} g_{3,k} g_{n-1} g_{2,j}) \\
 (14.5) \quad &= \tau(w_2 g_{n-1} g_{n-2} g_{n-1} g_{3,k} g_{2,j}) = \tau(w_2 g_{n-2} g_{n-1} g_{n-2} g_{3,k} g_{2,j}) \\
 &= s\tau(w_2 g_{n-2}^2 g_{3,k} g_{2,j}) \\
 &= s(t-1)\tau(w_2 g_{n-2} g_{3,k} g_{2,j}) + st\tau(w_2 g_{3,k} g_{2,j}).
 \end{aligned}$$

To analyze the right-hand side of (14.4), there are two cases:
Case 1: $j = 1$. Then

$$\tau(w_1 g_{n-1}^2) = ((t-1)s + t)\tau(w_1) = ((t-1)s + t)s\tau(w_2 g_{3,k})$$

while (14.5) yields in this case

$$\begin{aligned}
 \tau(g_{n-1} w_1 g_{n-1}) &= s(t-1)\tau(w_2 g_{n-2} g_{3,k}) + st\tau(w_2 g_{3,k}) \\
 &= (s^2(t-1) + st)\tau(w_2 g_{3,k}).
 \end{aligned}$$

Case 2: $j > 1$. Then

$$\begin{aligned}
 \tau(w_1 g_{1,j} g_{n-1}) &= \tau(w_1 g_{n-1} g_{n-2} g_{n-1} g_{3,j}) \\
 &= \tau(w_1 g_{n-2} g_{n-1} g_{n-2} g_{3,j}) = s\tau(w_1 g_{n-2}^2 g_{3,j}) \\
 &= s(t-1)\tau(w_1 g_{n-2} g_{3,j}) + st\tau(w_1 g_{3,j}) \\
 &= s(t-1)\tau(w_1 g_{2,j}) + st\tau(w_1 g_{3,j}),
 \end{aligned}$$

while (14.5) and induction yield in this case

$$\begin{aligned}
 \tau(g_{n-1} w_1 g_{1,j}) &= s(t-1)\tau(w_2 g_{n-2} g_{3,k} g_{2,j}) + st\tau(w_2 g_{3,k} g_{2,j}) \\
 &= s(t-1)\tau(w_1 g_{2,j}) + s^2 t\tau(w_2 g_{3,k} g_{3,j}) \\
 &= s(t-1)\tau(w_1 g_{2,j}) + st\tau(w_2 g_{2,k} g_{3,j}) \\
 &= s(t-1)\tau(w_1 g_{2,j}) + st\tau(w_1 g_{3,j}).
 \end{aligned}$$

We have now verified (14.3) in all cases, and hence τ satisfies (ii). $\square$

The main interest in τ stems from the fact that it can be normalized to yield a link invariant. In order to understand this, we first note that the generators g_i of H_n are units. In fact, $g_i(g_i + (1-t)) = t$ so we see that $g_i^{-1} = t^{-1}g_i + t^{-1} - 1$, and $\tau(g_i^{-1}) = t^{-1}s + t^{-1} - 1$. Now if we define θ by the equation $\tau(\theta g_i) = \tau(\theta^{-1} g_i^{-1})$, then

$$\theta^2 = \frac{\tau(g_i^{-1})}{\tau(g_i)} = \frac{t^{-1}s + t^{-1} - 1}{s} = \frac{s - t + 1}{st}$$

so θ lies in a quadratic extension of $\mathbf{Q}(s, t)$. Let $\hat{g}_i = \theta g_i$. Then the $\hat{g}_i$ satisfy the first two relations of (13.1) and therefore there is a homomorphism

$\pi_n : B_n \rightarrow H_n$ with $\pi_n(\sigma_i) = \hat{g}_i$, where B_n is the Artin braid group. Moreover,

$$\tau(\pi_n(\sigma_i^{-1})) = \tau(\pi_n(\sigma_i)) = \theta s$$

for all i , and

$$\tau(\pi_{n+1}(\alpha\sigma_n)) = s\theta\tau(\pi_n(\alpha))$$

for all $\alpha \in B_n$. It follows that if we set $\hat{\tau}(\alpha) = (\theta s)^{n-1} \tau(\pi_n(\alpha))$ for any n -string braid $\alpha \in B_n$, then

- (i) $\hat{\tau}(x^{-1}\alpha x) = \hat{\tau}(\alpha)$ for any $x \in B_n$, and
- (ii) $\hat{\tau}(\alpha\sigma_n^{\pm 1}) = \hat{\tau}(\alpha)$.

The above relations say that $\hat{\tau}$ is invariant under the so-called ‘‘Markov moves’’ which, by a theorem of Markov [1] implies that $\hat{\tau}(\alpha)$ depends only on the link $\hat{\alpha}$ obtained by joining the ends of the braid strings in $\mathbf{R}^3$.

Notice that if w is a word in the g_i of exponent sum $\varepsilon(w)$, then

$$\hat{\tau}(w) = (\theta s)^{n-1} \theta^{\varepsilon(w)} \tau(w).$$

It can be shown that, after a suitable change of variables, $\hat{\tau}$ is actually a Laurent polynomial in two variables.

Example. The trefoil can be obtained by joining the ends of the braid σ_1^3 . Thus (with $n = 2$), we have

$$\hat{\tau}(g_1^3) = (\theta s)\theta^3\tau((t-1)g_1^2 + tg_1) = \theta^4 s^3[(t-1)^2 s + (t-1)t + ts].$$

For the remainder of this section, we will essentially follow Springer [9]. Let $F = \mathbf{Q}(s, t)$ be the field of rational functions in two variables. Then τ_n is an F -linear functional on $FH_n = H_n \otimes F$. By (13.6), F is a splitting field for H_n , and by abuse of notation we will continue to denote by ξ_π the irreducible characters of FH_n . It is not difficult to see that any linear functional t on a complete matrix ring which satisfies $t(xy) = t(yx)$ must be a multiple of the trace by letting y range over the matrix units. This implies that there exist $\alpha_\pi \in F$ such that

$$\tau_n = \sum_{|\pi|=n} \alpha_\pi \xi_\pi.$$

The interesting result here is that there is a homomorphism φ from the ring of symmetric functions to F such that $\varphi(s_\pi) = \alpha_\pi$ where the s_π are the Schur functions (see §10). This allows us to express the α_π in terms of the $\alpha_{(p)}$ where (p) is the partition of p with one part. We then obtain a simple recursion formula for $\alpha_{(p)}$.

(14.6) *Let $\alpha_{(0)} = 1$ and $\alpha_{(k)} = 0$ for $k < 0$. Then for any partition $\pi = (\pi_1, \dots, \pi_r)$ we have*

$$\alpha_\pi = \det[\alpha_{(\pi_j + i - j)}]_{(1 \leq i, j \leq r)}.$$

PROOF. Identify $S^k \times S^{n-k}$ with the obvious Young subgroup of S^n . Let $\sigma \in S^k$ and $\rho \in S^{n-k}$. Then the desired formula is essentially a consequence

of

$$(14.7) \quad \tau_n(w_{\sigma\rho}) = \tau_k(w_\sigma)\tau_{n-k}(w_\rho),$$

which can be easily verified using (13.7), induction on $n - k$, and the definition of τ . This allows us to restrict τ_n to $H_{n,k}$ and use (13.7) to obtain

$$\sum_{\pi, \lambda, \mu} \alpha_\pi c_{\lambda\mu}^\pi \xi_{\lambda\mu}(w_{\sigma\rho}) = \sum_{\lambda, \mu} \alpha_\lambda \alpha_\mu \xi_\lambda(w_\sigma) \xi_\mu(w_\rho) = \sum_{\lambda, \mu} \alpha_{\lambda\mu} \xi_{\lambda\mu}(w_{\sigma\rho}),$$

where the sums range over partitions λ of k , μ of $n - k$, and π of n , and the $c_{\lambda\mu}^\pi$ are the Littlewood-Richardson coefficients. Since the $w_{\sigma\rho}$ are a basis for $H_{n,k}$ and the irreducible characters of $H_{n,k}$ are linearly independent, we can equate coefficients to obtain

$$(14.8) \quad \alpha_\lambda \alpha_\mu = \sum_{\pi} c_{\lambda\mu}^\pi \alpha_\pi$$

for all partitions λ of k and μ of $n - k$. However, it follows from (11.2) and (11.4) that

$$s_\lambda s_\mu = \sum_{\pi} c_{\lambda\mu}^\pi s_\pi$$

where s_π is the Schur function of type π . Since the Schur functions are a $\mathbf{Z}$ -basis for the ring of symmetric functions by (10.6), there is a ring homomorphism $\varphi : \Lambda \rightarrow F$ with $\varphi(s_\pi) = \alpha_\pi$. If we first use (10.5) with $\lambda = (p)$ to get $s_{(p)} = h_p$, then (10.5) becomes

$$s_\pi = \det[s_{(\pi_j + i - j)}]_{(1 \leq i, j \leq r)}$$

and the result follows by applying φ . $\square$

It remains to compute $\alpha_{(p)}$ for all p . For this purpose, the following lemma is useful.

(14.9) *Let $e_n = \sum_{\sigma \in S^n} w_\sigma$. Then $e_n g_i = t e_n$ for $1 \leq i < n$.*

PROOF. We proceed, as usual, by induction on n . Let $\rho_n = \sum_{i=0}^{n-1} g_{1,i}$. Then $e_n = e_{n-1} \rho_n$ by the definition of standard words. Furthermore, $\rho_n = 1 + g_{n-1} \rho_{n-1}$ and $\rho_{n-1} = 1 + g_{n-2} \rho_{n-2}$. Then

$$\begin{aligned} g_{n-1} \rho_{n-1} g_{n-1} &= g_{n-1}^2 + g_{n-1} g_{n-2} g_{n-1} \rho_{n-2} \\ &= (t-1) g_{n-1} + t + g_{n-2} g_{n-1} g_{n-2} \rho_{n-2}. \end{aligned}$$

Since $e_{n-1} g_{n-2} = t e_{n-1}$ by induction, we can left-multiply by e_{n-1} to get

$$\begin{aligned} e_{n-1} g_{n-1} \rho_{n-1} g_{n-1} &= (t-1) e_{n-1} g_{n-1} + t e_{n-1} + t e_{n-1} g_{n-1} g_{n-2} \rho_{n-2} \\ &= t e_{n-1} (1 + g_{n-1} + g_{n-1} g_{n-2} \rho_{n-2}) - e_{n-1} g_{n-1} \\ &= t e_{n-1} \rho_n - e_{n-1} g_{n-1} = t e_n - e_{n-1} g_{n-1}, \end{aligned}$$

whence

$$e_n g_{n-1} = (e_{n-1} + g_{n-1} \rho_{n-1}) g_{n-1} = t e_n.$$

For $i > 1$, write

$$\rho_n = \sum_{j < i-1} g_{1,j} + g_{1,i-1}(1 + g_{n-i}) + \sum_{j > i} g_{1,j}.$$

Then

$$\begin{aligned} \rho_n g_{n-i} &= \sum_{j < i-1} g_{1,j} g_{n-i} + g_{1,i-1}(g_{n-i} + g_{n-i}^2) + \sum_{j > i} g_{1,j} g_{n-i} \\ &= g_{n-i} \sum_{j < i-1} g_{1,j} + t g_{1,i-1}(1 + g_{n-i}) + g_{n-i-1} \sum_{j > i} g_{1,j}. \end{aligned}$$

By induction, we have $e_{n-1} g_{n-i} = e_{n-1} g_{n-i-1} = t e_{n-1}$, whence

$$e_n g_{n-i} = e_{n-1} \rho_n g_{n-1} = t e_{n-1} \rho_n = t e_n. \quad \square$$

From (14.9) we see that $e_n H_n$ is a one-dimensional right ideal affording the “principal character” $\xi_{(n)}$, namely $e_n \alpha = \xi_{(n)}(\alpha) e_n$ for all $\alpha \in H_n$, where ξ_n is the homomorphism $H_n \rightarrow \mathbf{Q}[t]$ defined by $\xi_{(n)}(g_i) = t$ for all i .

In particular, $e_{n-1} \rho_{n-1} = (1 + t + t^2 + \cdots + t^{n-2}) e_{n-1}$, from which it follows that

$$\begin{aligned} \tau(e_n) &= \tau(e_{n-1} + e_{n-1} g_{n-1} \rho_{n-1}) = \tau(e_{n-1}) + s \tau(e_{n-1} \rho_{n-1}) \\ &= \tau(e_{n-1}) \left[1 + s \frac{1 - t^{n-1}}{1 - t} \right]. \end{aligned}$$

Applying the above result to $\tau(e_i)$ for $i = n-1, n-2, \dots, 1$ and using $\tau(e_1) = 1$, we obtain

$$\tau(e_n) = \prod_{i=1}^{n-1} \left[1 + s \frac{1 - t^i}{1 - t} \right].$$

On the other hand, it follows from the results of §1 that $\xi_\pi(e_n) = 0$ for $\pi \neq (n)$, hence $\tau(e_n) = \alpha_{(n)} \xi_{(n)}(e_n)$. Furthermore, it is clear that $e_n w_\sigma = t^{l(\sigma)}$ for some integer $l(\sigma)$ which is usually called the “length” of σ , and if we put $p_n(t) = \sum_{\sigma \in S^n} t^{l(\sigma)}$ then $\xi_{(n)}(e_n) = p_n(t)$. Since

$$e_n = e_{n-1} \rho_n = \cdots = \prod_{i=1}^n \rho_i$$

we obtain the well-known formula

$$p_n(t) = \sum_{\sigma \in S^n} t^{l(\sigma)} = \xi_{(n)}(e_n) = \prod_{i=1}^n \xi_{(n)}(\rho_i) = \prod_{i=1}^n \frac{1 - t^i}{1 - t}.$$

Finally, since $\tau(e_n) = \alpha_{(n)} \xi_{(n)}(e_n) = \alpha_{(n)} p_n(t)$, we have proved

(14.10) For any integer $n \geq 1$ we have

$$\alpha_{(n)} = \frac{1}{p_n(t)} \prod_{i=1}^{n-1} \left[1 + s \frac{1 - t^i}{1 - t} \right]. \quad \square$$

Bibliography

1. J. Birman, *Braids, links, and mapping class groups*, Ann. of Math. Stud., no. 82, Princeton Univ. Press, Princeton, NJ, 1974.
2. N. Bourbaki, *Groupes et algebres de Lie*, IV, V, VI, Masson, Paris, 1982.
3. P. Freyd, J. Hoste, W. Lickorish, K. Millet, A. Ocneanu and D. Yetter, *A new polynomial invariant of knots and links*, Bull. Amer. Math. Soc. (N.S.) **12** (1985), 183–312.
4. G. James, *The representation theory of the symmetric groups*, Lecture Notes in Math., vol. 682, Springer-Verlag, New York, 1978.
5. V. F. R. Jones, *Hecke algebra representations of braid groups and link polynomials*, Ann. of Math. (2) **126** (1987), 335–388.
6. D. Kazdan and G. Lusztig, *Representations of Coxeter groups and Hecke algebras*, Invent. Math. **53** (1979), 165–184.
7. S. Lang, *Algebra*, Addison-Wesley, New York, 1978.
8. I. G. Macdonald, *Symmetric functions and Hall polynomials*, Oxford Univ. Press, Oxford, 1979.
9. T. A. Springer, *Jones' trace function on the Hecke algebra of symmetric groups*, Topics in Algebra, Banach Center Publ., vol. 26, part 2, PWN-Polish Scientific Publishers, Warsaw, 1990.

